DOCUMENT RESUME

ED 390 268 FL 023 460

TITLE	Report of the Task Force on Testing Standards (TFTS) to the International Language Testing Association (ILTA).
INSTITUTION	International Language Testing Association.
PUB DATE	Sep 95
NOTE	186p.
AVAILABLE FROM	International Language Testing Association, NLLIA Language Testing Research Centre, Department of Applied Linguistics and Language Studies, University of Melbourne, Parkville, Victoria 3052, Australia.
PUB TYPE	Reports - Descriptive (141) -- Reference Materials - Directories/Catalogs (132)
EDRS PRICE	MF01/PC08 Plus Postage.
DESCRIPTORS	*Academic Standards; Bibliographies; Definitions; *Foreign Countries; *Language Tests; Standards; Surveys; *Testing

ABSTRACT

The Task Force on Testing Standards (TFTS) of the International Language Testing Association was charged to produce a report of an international survey of language assessment standards, to provide for exchange of information on standards and for development of a code of practice. Contact with individuals in both language testing and the broader educational testing domain in countries around the world resulted in the collection of 110 documents on standards. These documents are described here, in bibliographic format. Actual copies of the documents reside with TFTS. An introductory section details the survey's methodology and summarizes its results. In a brief conclusion, four specific recommendations are made. (1) reconstitution of the TFTS; (2) creation of a summary of standards information for each country; (3) active pursuit of world standards in language testing, beginning with a definition of the term "standard"; and (4) establishment of the bibliography as a dynamic and ongoing database. Appended materials, the bulk of the document, include: the TFTS charge and inquiry letters; summaries of materials received from each country responding; and further general comments and notes on contacts made. Two chapters of a published text on test construction and evaluation are available in the first 100 printed reports only and are not available in the ERIC document or any other copies of this report. (MSE)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 390 268

Report of the Task Force on Testing Standards (TFTS) to the International Language Testing Association (ILTA)

by

the ILTA TFTS:

Fred Davidson, University of Illinois, USA (TFTS Chair)
J.Charles Alderson, Lancaster University, UK
Dan Douglas, Iowa State University, USA
Ari Huhta, University of Jyväskylä, Finland
Carolyn Turner, McGill University, Canada
Elaine Wylie, Griffith University, Australia

Full addresses of the TFTS members
are given in Appendix Two, below.

September, 1995

PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

- This document has been reproduced as received from the person or organization originating it.
- Minor changes have been made to improve reproduction quality.
- Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

FL 023 460

BEST COPY AVAILABLE

Copyright © 1995, International Language Testing Association

The material in Appendix Five is copyrighted, 1995, by Cambridge University Press and reproduced by permission of Cambridge University Press. This statement applies only to the first one hundred printed copies of this report. All other copies of this report, per agreement with C.U.P., do not contain Appendix Five.

ACKNOWLEDGMENTS

The TFTS would like to thank many people who made this project feasible. First, the members, officers and executive board of ILTA deserve special mention. Throughout the time we worked, our colleagues in ILTA provided very valuable feedback and suggestions about our activities. Special thanks are also due to Bonny Peirce, who was extremely helpful in securing information from South Africa. We also acknowledge the support of the University of Illinois Computing and Communications Office, and in particular Mark Zinzow, for establishing an e-mail discussion group via which we conducted our business, and for arranging the "ftp" and e-mail pick-up of this document for review at the 1995 ILTA Annual General Meeting at LTRC 1995. We also wish to thank Doug Mills of the Division of English as an International Language, who helped to place a copy of this report on the World Wide Web for the same purpose. And to all those who attended the AGM and provided feedback, many thanks. Finally, we thank our respective families (sine qua non) for bearing with us during the assembly of this material and completion of this project.

FOR FURTHER INFORMATION

For further information on this report, or to order additional copies, please contact ILTA at the following address.

Dr. Tim McNamara
Secretary, ILTA
NLLIA Language Testing Research Centre
Department of Applied Linguistics and Language Studies
The University of Melbourne
Parkville, VIC 3052, Australia
tel. +61-3-344-4207
fax. +61-3-344-5163
e-mail: tim_mcnamara@muwayf.unimelb.edu.au

At press time, ILTA plans to place a copy of this report in the ERIC microfiche system and possibly on the World Wide Web [WWW]. The ERIC and WWW copies will not contain Appendix Five, as per an arrangement with Cambridge University Press.

Table of Contents

EXECUTIVE SUMMARY----------8

INTRODUCTION----------9

METHODOLOGY----------9

RESULTS AND DISCUSSION----------11

Structure of the TFTSTB ----------11

TFTSTB Records: Summary Tables ----------13

Discussion ----------16

CONCLUSION----------18

APPENDIX ONE: ILTA TFTS CHARGE LETTER----------19

APPENDIX TWO: TFTS ENQUIRY LETTER: MASTER COPY/TEMPLATE----------20

APPENDIX THREE: THE TFTS TEXTBASE [TFTSTB]----------23

AU01: Australia: "Ethical Considerations in Ed. Testing..." -------23

AU02: Australia: "Assessment, referral and placement. No 17..." ---24

AU03: Australia: "School Certificate Grading System: Course..." ---25

AU04: Australia: "Subject Manual 5A: Languages other than ..." ----26

AU05: Australia: "Educational objectives being tested in the..." --27

AU06: Australia: "Australian Scholastic Aptitude Test ..." --------28

AU07: Australia: "NAFLaSSL Information Manual, 1993" --------------30

AU08: Australia: "ACCESS Test Specifications..." ------------------31

AU09: Australia: "Occupational English Test for Overseas..." ------32

AU10: Australia: "NLLIA Japanese for Tourism and Hospitality..." --33

AU11: Australia: "Direct testing of general proficiency ..." ------34

AU12: Australia: "Discussion Papers 1-21 (1986/7/8)..." -----------35

AU13: Australia: "Tasmanian Certificate of Education..." ----------40

AU14: Australia: [Three documents from the Sec. Ed. Authority] ----41
AU15: Australia: "The Process of Assessment, Grading, and ..." ----44
AU16: Australia: "Faculty Assessment Policies" [Griffith Univ.] ---45
AU17: Australia: "ESL Development: Language and Literacy in..." ---47
AU18: Australia: "Ethical Guidelines..." [Migrant Ed.] ----------49
Explanatory Notes regarding records from Canada ----------51
CA01: Canada: "Principles for Fair Student Asmt. Practices..." ----53
CA02: Canada: [CPA] "Guidelines for Ed. and Psych. Testing ..." ---55
CA03: Canada: "Annotated List of French Tests" ----------57
CA04: Canada: "ESL Instruction in the Junior High School ..." -----58
CA05: Canada: "General Policy for Educational Evaluation..." ------59
CA06: Canada: "[Guide to classroom eval.: sec. language...]" ------60
CA07: Canada: "[Guide to classroom eval.: formative...]" ----------61
CA08: Canada: "[Developing a criterion-referenced...]" ----------61
CA09: Canada: "Evaluation, Module 4." ----------62
CA10: Canada: "[Definition of domain for French as a Sec...]" -----63
CA11: Canada: "[The effects of the language used in eval. ...]" ---63
CA12: Canada: "[Guide to test construction, ESL, 2nd Cycle...]" ---64
CA13: Canada: "[Collection of ESL reading test items/tasks...]" ---65
CA14: Canada: "[Guide to evaluating speaking in class, ESL,...]" --66
CA15: Canada: "[Guide to test construction, ESL, 2nd Cycle...]" ---67
CA16: Canada: "[Guide to test construction, ESL, 1st Cycle...]" ---67
CA17: Canada: "The Ontario Test of Engl. as a 2nd Lang. ..." ------68
CA18: Canada: "The Canadian Test of Engl... (CanTEST)..." ---------69
CA19: Canada: "English Language Program: Intens. Curr. ..." -------71
CA20: Canada: "Carleton Acad. Engl. Lang. Asmt. (CAEL)..." --------72
CH01: China: "A Brief Introduction to the Engl. Lang. Exam..." ----74

EU01: Europe/Int'l: "Assoc. Lang. Testers in Europe (ALTE)..." ----74
EU02: Europe/Int'l: "International Baccalaureate Exams..." --------75
FI01: Finland: "[National certificate: Finnish lang. test...]" ----75
FI02: Finland: "[National certificate: test specifications...]" ---78
FI03: Finland: "The Finnish Matriculation Examination" ----------80
FR01: France: "DELF (Diplôme d'Etudes en Langue Française)..." ----81
FR02: France: "1994. Guide du concepteur de sujets. DELF-DALF" ----82
FR03: France: "Monitoring Education-For-All Goals..." ----------84
GE01: Germany: "[Administering the Goethe-Institut tests...]" -----86
GE02: Germany: "[Lesser German Language Diploma. Greater German...]"88
HK01: Hong Kong: "The Work of the H.K. Exam. Authority..." --------90
HK02: Hong Kong: "An Introduction to Educational Assessment..." ---91
HK03: Hong Kong: "General Introduction to Targets..." ----------92
HK04: Hong Kong: "Public Examinations in H.K., 1993" ----------93
HK05: Hong Kong: "Statistics Used in Public Examinations..." ------93
IN01: India: "Handbook of Evaluation in English..." ----------94
IR01: Ireland: [NCCA] "Junior Certificate..., Leaflet..." ---------95
MA01: Mauritius: "The Certificate of Primary Education..." --------96
NA01: Namibia: "Draft Manual of Standards in English..." ----------97
NE01: The Netherlands: "Certificate of Dutch as a Foreign..." -----98
NE02: The Netherlands/Int'l: [IAEA] [Various material] ----------99
NE03: The Netherlands/Int'l: [IAEA] "Standards for Design..." ----104
NZ01: New Zealand: "Regulations and Prescriptions Handbook..." ---106
PO01: Portugal: [Materials to support teachers nationwide...] ----107
SA01: South Africa: "Handbook for English: GEC Exam..." ----------107
SA02: South Africa: "User Guide 1, General Handbook: Adult..." ---109
SA03: South Africa: "Standards -- The Loaded Term..." ----------110

SD01: Sweden: "[Central examinations in the senior ...]" ---------110
SE01: Seychelles: "The Certificate of Proficiency in Engl..." ----113
SI01: Singapore: "PSLE Information..., Assessment Guide..." ------114
SW01: Switzerland: "Language B Guide, First Edition, 1994." ------114
SW02: Switzerland: "Pedagogic policy..., Guidelines for..." ------115
TA01: Tanzania: "Continuous Assessment: Guidelines..." -----------116
UG01: Uganda: "Language Examinations, Ordinary and Adv..." -------117
UK01: England: "Handbook for Centres: All Cambridge Exams..." ----117
UK02: England: "The Common Syllabuses at Levels A, B, ..." -------119
UK03: England: "Issues in Public Examinations (1991)" ------------120
UK04: Fngland: "Examinations: Comparative and International..." --121
UK05: England: "GCSE Mandatory Code of Practice..." --------------121
UK06: England: [various material from Univ. London] --------------122
UK07: England: [LCCI] "Centre Application Form & Fee Sheet" ------123
UK08: England: "Introduction to the National Language Stds..." ---124
UK09: Wales: "Syllabuses for French, German, and Span..." --------125
UK10: England: "The BPS Statement and Certificate..." ------------125
UK11: England: [BPS] "Psychological Testing: A Guide" ------------127
UK12: England: "The International Encyclopedia of Ed. Eval." -----127
UK13: England: [BPS] "Code of Conduct Ethical Principles..." -----129
UK14: England: "Psychological Testing: A Practical Guide..." -----130
UK15: England: "Foreign Language Testing. Specialised..." -------130
UK16: England: "Foreign Language Testing Supplement..." ----------131
UK17: England: "Testing Bibliography, Nov 1994. CILT" -----------131
UK18: England: "Info. Sheet 10: Guide to GCSE 'A' Level..." ------132
UK19: England and Wales: "GCE A and AS Code of Practice..." ------132
UK20: England and Wales: "Modern Foreign Languages..." -----------133

UK21: England: "Certificates in Commun. Skills in English ..." ---134

UK22: England: "The IELTS Specifications..." ----------136

UK23: Scotland: [Brochures:] "National Certificate and You..." ---137

UK24: Scotland: [Scot. Exam. Board: Various documents] ----------138

UK25: Scotland: "Communicative Language Testing: a Resource..." --139

UK26: Scotland: "Curriculum and Assessment in Scotland..." -------139

UK27: Scotland: "Assessment 5-14. Improving the Quality..." -----141

UK28: Scotland: "Taking a Closer Look. A Resource Pack..." ------141

UK29: Scotland: [Scot. Exam. Board, Cmte. on Testing: various] ---142

US01: U.S.A.: [APA/AERA/NCME] "Standards for Ed. & Psych. ..." ---143

US02: U.S.A.: "ETS Standards for Quality and Fairness" ----------153

US03: U.S.A.: "Mental Measurement Yearbook" ----------155

US05: U.S.A.: "Questions to Ask When Evaluating Tests." ----------157

US06: U.S.A.: "Criteria for Evaluation of Student Assess..." -----157

US07: U.S.A.: "Principles and Indicators for Student Assess..." --158

US08: U.S.A.: "Implementing Perform. Assessments: A Guide..." ----160

APPENDIX FOUR: CONTACT NOTES----------163

APPENDIX FIVE: CHAPTERS ONE AND 11 OF ALDERSON, CLAPHAM AND WALL (1995) LANGUAGE TEST CONSTRUCTION AND EVALUATION (CAMBRIDGE U.P.) [PER ARRANGEMENT WITH THE PUBLISHER, AVAILABLE IN THE FIRST ONE HUNDRED PRINTED REPORTS ONLY, AND NOT IN THE ERIC DOCUMENT OR ANY OTHER COPIES OF THIS REPORT.]----------184

EXECUTIVE SUMMARY

-- Pursuant to an October 1993 charge given by Charles Stansfield, the ILTA president (pro-tem), the Task Force on Testing Standards was constituted and in fulfillment of its charge has produced a report of a international survey of assessment standards.

The purpose of this project was twofold: first, it provides a general resource for scholars of educational and language assessment standards, and second, it serves as a specific resource for later ILTA efforts to draft its own code of practice, in that ILTA can consult this report to obtain information about extant documents of that nature.

-- This was accomplished by contact with colleagues in the educational assessment community, both language testing and the broader educational testing domain.

-- 110 respondents sent documents on standards to TFTS members. These documents are described in a bibliographic textual database, enclosed in Appendix Three below.

-- Great variety was observed in the respondents' definition of "standard" as well as other variables of interest.

-- For that reason, if ILTA re-constitutes the TFTS, we recommend that its first action should be to agree on a definition of 'standard'. Additional recommendations following on from this Report appear in the Results/Discussion and Conclusion, below.

Introduction

In July of 1993, Charles Stansfield, the pro-tem president of the newly-formed International Language Testing Association (ILTA) proposed that an ILTA Task Force on Test Standards (TFTS) be formed. After discussion in ILTA, that group was formed in October of 1993. The mission of the TFTS would be to survey world standards of educational measurement, with specific reference to language assessment. The following persons were named to be members of the TFTS:

J.Charles Alderson, Lancaster University, UK
Fred Davidson, University of Illinois, USA (TFTS Chair)
Dan Douglas, Iowa State University, USA
Ari Huhta, University of Jyväskylä, Finland
Carolyn Turner, McGill University, Canada
Elaine Wylie, Griffith University, Australia

This report, jointly authored by the TFTS, is the product of work since the date of the TFTS charge (see Appendix One for a copy of the charge letter). The purpose of this project was twofold: first, it provides a general resource for scholars of educational and language assessment standards, and second, it serves as a specific resource for ILTA's move forward to draft its own code of practice, in that ILTA can consult this report to obtain information about extant documents of that nature.

With the submission of this report to ILTA, the duties of the TFTS are formally discharged and it is disbanded.[1]

Methodology

To effect the charge given it by ILTA, the TFTS chose to liaise with colleagues and draft an 'enquiry letter' (see Appendix Two). This letter states the basic intent of ILTA and TFTS in this project. The letter was sent to a vast array of agencies involved in educational assessment around the world. The agencies which received the letter were the product of brainstorming among TFTS members in consultation with ILTA, via, for example, the ILTA Annual Business Meeting at the 1994 Language Testing

[1] As of the date of this report, the materials collected by the TFTS are in the keeping of individual TFTS members. We anticipate ongoing discussions in ILTA about whether that material should be centralized in some library and made available through inter-library loan. Members of the present ILTA TFTS remain available to help facilitate that process, but as of completion of this report, we see our charge as essentially fulfilled.

Research Colloquium in Washington, DC, USA.

While the letter itself did spark a great deal of response, the TFTS also used personal contact as a means to obtain material to discharge its duties. These contacts include phone and e-mail communication and various personal communication in the world community of educational measurement specialists.

The TFTS coordinated its activities largely by e-mail, using a LISTSERVer (e-mail discussion group) established at the University of Illinois. In addition, members of the Task Force discussed progress at any face-to-face opportunity, for example at professional meetings. Finally, critical information was disseminated by letters, for example notification of acceptance of a TFTS report at the TESOL 1995 meeting in Long Beach.

The TFTS information gathering produced a collection of documents. Some of these documents are short reports, others are complete books. Each shares the common feature of having been published or used as an in-house document by the source providing the document. The majority of the documents were obtained by the enquiry letter, phone call, or discussion with colleagues.

The TFTS assembled its findings into two forms. First, actual copies of the documents reside with the TFTS. Second, a collection of bibliographic information about the documents has been assembled. That collection, called the TFTS 'TextBase' ('TFTSTB') appears in its entirety in Appendix Three. An entry in TFTSTB is formed based on a return reply from a source contacted in the manner described above. The records in the TFTSTB are usually reflective of something provided by a source whom we contacted. If the source returned one document, then that single document constitutes the TFTSTB entry. If the source returned several documents, or some sort of document series, the TFTSTB entry is actually a reference to multiple documents. In all instances, we tried to let whatever the source sent us stand as a guideline for separability into a unique record in the TFTSTB.

Appendix Four below presents two types of information. First, TFTS members reflect on the experience of collecting this data and share some summary comments about findings for particular sets of records. Second, TFTS members share notes on persons they contacted who did not provide documents for inclusion in the TFTSTB. This non-response information is presented for the historical record only.

Finally, in Appendix Five, we provide a copy of Chapters One and 11 of Alderson, Clapham and Wall (1995). Chapter One contains information that explains the data reported in Chapter 11, which deals with topics that are very similar to and hence quite relevant to the mission of the TFTS. In addition, as can be seen, that book had influence on the entire TFTS Report in the organizational rubric for commentary in TFTS records.

RESULTS AND DISCUSSION

Structure of the TFTSTB

The structure of the TFTSTB is itself a result of extensive discussion of TFTS members, largely by e-mail. Additionally, it reflects the actual nature of the data received from sources. Hence, the TFTSTB structure is more of a 'result' than an element of our methodology.

The various information in each TFTS record is an attempt to capture as much detail as possible about each document without having the document itself. The TFTSTB is not only a bibliographic database. It is also a record of scholarly reaction to the material we received.

Each record in the TFTSTB has the following fields:

Record: Arbitrary sequential number within a two-character country code, e.g. 'UK01' is the first record for the United Kingdom, 'CA02' is the second record for Canada, etc. There is no meaning to the sequence of records within a given country -- that ordering is also arbitrary, and derives from the timing of responses to the enquiry letter.

TFTSmem: Name of TFTS member(s) who prepared the information for this record. Holder(s) of a copy of the document cited in the record, unless the item is widely-available in libraries (e.g. record US03).

Country: of authorship of the document(s) referenced in this record

Auth/Pub: of the document(s). Name(s) of authors (if any) and publisher, if any; otherwise, name of organizational author. If the author and publisher differ, that is clarified in the entry.

Title(s): of the document(s). Precise title(s)have been given wherever they were provided. Otherwise descriptive. Date(s) of publication are also given in this field.

Language: in which the document is written

Contact: Address, phone/e-mail/fax numbers etc. to obtain copies of and/or information about the document(s)

Govt/Priv: 'govt' = the document(s) is/are primarily or originally sponsored and/or authored by a government agency. If the document(s) are assumed to be published in the private sector, and 'priv' appears. Finally, some documents are

a 'collaboration' between government and private interests.

Lang/Ed: 'lang' = the document(s) is/are primarily concerned with language testing in particular. Otherwise the document is assumed to deal with the broader domain of general educational and psychological human measurement, and 'Ed' appears.

StanDef: A coding for the primary definition of "standard" used by or implied by the document(s). Documents may incorporate or refer to several interpretation of "standard", and in that case, the most prevalent interpretation was selected. Coding was done by the TFTS Chair in consultation with the person's name appearing under 'TFTSmem'. Specific codes for StanDef are:

'guideline' if "standard" refers to a guideline of good practice; textbooks and materials which resemble textbooks are included in this code

'other' if it is not clear how the supplier of the document interpreted the term "standard"; i.e., if it is not clear why this document was submitted in response to the TFTS request letter, or if the entry refers to some definition of "standard" other than guideline, performance or test.

'performance' if "standard" refers to a performance criterion, e.g. the standard of being able to fly an airplane or negotiate a business transaction, or "standard" refers to a particular ability level or levels, or "standard" refers to a cutscore or cutscores on a distribution

'test' if "standard" refers to a test or tests described and/or reviewed by the document(s), or if "standard" refers to some actual test(s) and/or test(s') manual(s).

Comments: about the document, written by 'TFTSmem'

Figure 1:
TFTSTB: Record Structure

In addition to the fields explained in Figure 1, a header appears at the top of each record, containing the record number, country, and an abbreviated title. This is used to generate the table of contents for the report and may be ignored.)

Within the 'Comments' field appears extensive information and scholarly comment, often using the following subfields (adapted from Alderson, Clapham and Wall (1995) with adaptation by Turner[1]).

[1] Alderson, J.C., C. Clapham and D. Wall. 1995. Language Test Construction and Evaluation. Cambridge University Press..

Objective/Purpose:
Target group/Audience intended:
Procedure:
Scope of influence:
Summary:
Comment/Commentary: (this is a subfield of 'Comments')

Figure 2:
TFTSTB: Subfields Under 'Comments'

Within the TFTSTB, there is structural variability for the sub-fields under 'Comments'. Some records are for multiple documents, e.g. the first IAEA entry (record NE02), and necessarily follow a different format. In other cases, the TFTS member(s) who prepared the record's information chose to organize the comments in a different manner, perhaps lengthening one sub-field (e.g. 'Commentary' for the APA/AERA/NCME Standards, record US01). The TFTS does not treat the 'Comments' field as a strict organizational tool, but rather views it as an optional layout. In that way, we maximize the individual contribution of each Task Force member to this report, and heighten the scholarly reactive nature of the TFTSTB. The evolution of the 'Comments' field was also a result of extensive TFTS discussion, as was its flexibility.

TFTSTB Records: Summary Tables

All tables given here were produced from analysis of a dataset containing the record number and the values for the three fields: 'govtpriv', 'langed', and 'standef'. The analyses were run in PC-SAS for Windows (Version 6.10, 1994). The tables given are cut-and-pasted from the SAS output listing.

COUNTRY		Frequency	Percent	Cumulative Frequency	Cumulative Percent
AU	Australia	18	16.4	18	16.4
CA	Canada	20	18.2	38	34.5
CH	China (PRC)	1	0.9	39	35.5
EU	Europe	2	1.8	41	37.3
FI	Finland	3	2.7	44	40.0
FR	France	3	2.7	47	42.7
GE	Germany	2	1.8	49	44.5
HK	Hong Kong	5	4.5	54	49.1
IN	India	1	0.9	55	50.0
IR	Ireland	1	0.9	56	50.9
MA	Mauritius	1	0.9	57	51.8
NA	Namibia	1	0.9	58	52.7

NE	Netherlands	3	2.7	61	55.5
NZ	New Zealand	1	0.9	62	56.4
PO	Portugal	1	0.9	63	57.3
SA	South Africa	3	2.7	66	60.0
SD	Sweden	1	0.9	67	60.9
SE	Seychelles	1	0.9	68	61.8
SI	Singapore	1	0.9	69	62.7
SW	Switzerland	2	1.8	71	64.5
TA	Tanzania	1	0.9	72	65.5
UG	Uganda	1	0.9	73	66.4
UK	United Kingdom	29	26.4	102	92.7
US	United States	8	7.3	110	100.0

Table 1:
Records per Country in the TFTSTB

The distribution of countries represented in the TFTSTB is a reflection of several variables. First, it derives from the makeup of the TFTS itself: the distribution of TFTS members worldwide conditioned whom they were able to contact. Second, it reflects the response rate -- members of the TFTS did contact a wider number of colleagues, but some did not respond (see Appendix Four).

GOVTPRIV	Frequency	Percent	Cumulative Frequency	Cumulative Percent
collaborative	2	1.8	2	1.8
government	43	39.1	45	40.9
other	3	2.7	48	43.6
private	62	56.4	110	100.0

Table 2:
Government / Private (GovtPriv) Sources

It is interesting to note the distribution of values in Table 2. Although 56% of the records are private, there are not many 'collaborative' or 'other' record sources. The TFTS seems largely to be a database of governmental and private sources.

LANGED	Frequency	Percent	Cumulative Frequency	Cumulative Percent
education	54	49.1	54	49.1
language	56	50.9	110	100.0

Table 3:
Domain: Language vs. Education

Table 3 shows an interesting even split. About as many

records are for material focused solely on language as those focused more broadly on educational and psychological measurement.

STANDEF	Frequency	Percent	Cumulative Frequency	Cumulative Percent
guideline	58	52.7	58	52.7
other	10	9.1	68	61.8
performance	15	13.6	83	75.5
test	27	24.5	110	100.0

Table 4:
Definition of Standard

The majority of records in the TFTSTB are for "standard" as it refers to a guideline for good practice. The second most common type of record is "standard" as it refers to a precise test or set of tests. The third most common definition of "standard" as a performance criterion, level, or cutscore. While we recognize the subjectivity involved in categorizing these documents, we note that the most common definition we see in our material is that of standard as guideline. Nonetheless, there is a distribution in Table 4, and a more detailed reading of, in particular, the comments fields of the TFTSTB records shows even greater substantive disagreement about the precise definition of "standard".

Discussion

There is great variety in the TFTS in interpretation of terms like "standard", "guideline", "code of practice" and the like. That finding is carried over in the commentary and notes of TFTS members given in Appendix Four. However, there are also some interesting points of discovery to be found in those notes. Based on those notes, the trends in the tables above, and the general discussions within the TFTS and between the TFTS and ILTA, we would like to interpret our findings a bit further.

First, there are clearly records in the TFTSTB which future ILTA standard-setting initiatives should consult. For example, Alderson says of FR03: "This [record] is interesting to ILTA because it represents an international attempt to reach agreement on the qualities of measuring instruments." There was also discussion among TFTS members and ILTA audiences (at TFTS presentations) that the APA/AERA/NCME document should be consulted closely as ILTA considers authoring its own code (see record US01). At the same time, FR03 -- and other discussions among TFTS members and ILTA audiences -- serve to remind us of the "I" in ILTA's name and the consequent importance of attending to various national and regional needs. We believe that measurement practice is a culturally-bound phenomenon, and ILTA's future work on standard-setting should reflect such boundedness. The process of writing a set of international standards is quite complex, and while certain that the TFTSTB records can serve as excellent thought stimuli to such an endeavor, we submit that no single TFTSTB record stands as a model of what ILTA should create.

In Appendix Four, Alderson also makes an interesting remark in his discussion of some correspondence from Patricia Broadfoot, which suggests "that our original request has been misinterpreted as pertaining exclusively to language testing." More generally, the definitional range of "standard" we have observed may be partly due to our methodology, and so we encourage ILTA to revisit the TFTSTB periodically and revise it. Such revision is also necessary because many of the sources are revised periodically, and because sources disappear and appear as educational systems change and evolve.

It is interesting to note, in regards to evolution of educational systems, that the TFTS members all used a combination of personal contact and unsolicited letter writing to obtain documents. As teaching systems change, the best way to obtain information about new resources is probably the very real network that also evolves among educators -- at conferences, in publications, and now, on the world computer networks. This suggests that the TFTSTB might become a dynamic ILTA entity -- possibly on a computer network -- which is updated regularly as

ILTA members participate in or discover documents relevant to its purposes.

Such networking is essential for a project like this, because the unsolicited letter-writing yielded a response rate which (as Huhta puts it in Appendix Four) "varied considerably". What is more, respondents sometimes reported back to TFTS members not by supplying documents but by stating that they felt they had nothing that matched our interests. Given the various interpretations of "standard" and "guideline", and given possible inadequacy of our methodology, it is possible that those respondents did in fact have something of interest to ILTA but mis-interpreted our interest. Networking can resolve this as well, for it can allow ILTA members to portray the range of ILTA's interests internationally -- the "I" in our name comes to have yet another influence: harvesting a wide variety of material. In the interest of such future growth, we have provided names of two types of non-respondents in Appendix Four: those from whom we got a response of the nature here described ("thanks for asking, but we don't have anything like that."), and those from whom we did not hear at all. We encourage readers of this report to let ILTA know of any contacts with these individuals, or with individuals not listed, and write the ILTA Secretariat.

By far the most pressing need for future clarity of international standards is to add additional countries and regions to the TFTSTB. The TFTS agree that country summaries are extremely valuable -- e.g. that authored by Turner (just before CA01) or the "Oz" remarks by Wylie in Appendix Four. We must emphasize again that the data we collected reflect our particular locations and networking strengths. **In not including any country or region in the TFTSTB, we do not wish to slight any individuals, any institutions, any governments, or any professional organizations**.

We submit that growth of the TFTSTB should proceed on several fronts: continued refinement of the present TFTSTB records, addition of data from non-respondents, and addition of data from countries and regions not reported here.

CONCLUSION

In closing, we would like to make some specific recommendations to ILTA:

(1) ILTA should consider re-constituting the TFTS, or a body like it, to continue and further formalize this line of work. (With the submission of this report, the work of this TFTS draws to a close, and per its mandate, it is disbanded.) Whether or not the ongoing work should be done by a special task force or by a standing committee is a matter for ILTA to determine.[3]

(2) As one task we recommend for the re-constituted TFTS would be to write 'country summaries' of the type which Turner volunteered to author for Canada (just before CA01 in the TFTSTB) or Wylie's "Oz" remarks in Appendix Four. The TFTS agree that this is a very valuable type of writeup.

(3) ILTA should pursue actively its role in setting world standards of language testing.

(3a) ILTA should first do so by agreeing on a definition of the term "standard". Perhaps this should be the first task of the re-constituted group.

(3b) The common definition could and should be informed by the content of the documents represented in the TFTSTB, as can be any standard-setting initiatives ILTA undertakes.

(4) Due to the potential growth and change of the TFTSTB, ILTA consider establishing it as an ongoing, dynamic (possibly computerized) entity, to which many ILTA members can contribute information.

That said, we -- the members of the TFTS -- speak as one voice to note that we enjoyed this scholarly exercise. We sense that ILTA is entering a world of great variety and flexibility as regards its role in standard-setting. This is a challenging and exciting academic direction for ILTA in the coming years.

---

[1] By the time this document reached press, ILTA already constituted such a group.

APPENDIX ONE: ILTA TFTS Charge Letter

[ILTA Letterhead]

July 27, 1993

[TFTS Member]
[Address]

Dear [TFTS member]

This constitutes your letter of appointment as a member of the Task Force on Test Standards of the International Language Testing Association (ILTA). The mandate of the task force is to collect information on test standards developed throughout the world and report to the ILTA membership on what can be learned from those standards documents.

As you know, many members of ILTA are interested in seeing the organization develop a set of voluntary standards for language tests that could be disseminated throughout the world. The work of the Task Force on Test Standards will provide background information on existing standards for tests in general that may be in effect throughout the world.

This information could be used by ILTA to gain an understanding of how standards documents are developed. It will also help ILTA develop a sensitivity to concept of good testing practice in different countries. Should ILTA decide to develop a set of standards, such information will provide guidance to the effort.

I have asked [the TFTS] to give an oral report on the findings at the 1994 annual business meeting of ILTA, which will be held in Washington, DC. A written report will be due in one year.

Thank you for your willingness to serve ILTA in this way.

Sincerely,

Charles W. Stansfield
ILTA President

APPENDIX TWO: TFTS Enquiry Letter: Master Copy/Template

(Date)

(Agency/Bureau/Person)
(Address)
(Address)
(Address)

Dear (etc.):

The new International Language Testing Association (ILTA) has formed a Task Force on Test Standards (TFTS). The mandate of the TFTS is to research the history and current status of educational assessment standards worldwide, both of assessment generally and with particular reference to language testing. By 'standards', we refer to documents, guidelines, policies, books and other materials which guide educators in construction or selection of educational assessment instruments. While textbooks of educational testing can serve to direct assessment procedures, we would include them in our request only if the texts clearly contain material on standard setting, e.g. a chapter or appendix on the topic.

We are forming a collection of such material in preparation of a written report to ILTA. To help us, we would like to know if you have any such material, and if so, how we might obtain copies.

We thank you in advance for any assistance you can offer.

Sincerely,

(TFTS member's name)
(TFTS member's title)

enclosure: ILTA brochure

ILTA TFTS:
---

(1) Fred Davidson, TFTS Chair
Division of English as an International Language (DEIL)
University of Illinois at Urbana-Champaign (UIUC)
3070 Foreign Languages Building (FLB)
707 S. Mathews
Urbana, IL 61801,
USA
tel: +217-333-1506
fax: +217-244-3050
e-mail: fgd@ux6.cso.uiuc.edu

(2) J.Charles Alderson
Dept. of Linguistics and Modern English Language
Lancaster University
Bowland College
Lancaster, LA1 4YT,
UK
tel: +01524-59-3029
fax: +01524-84-3085
e-mail: c.alderson@lancaster.ac.uk

(3) Dan Douglas
Department of English
316 Ross Hall
Iowa State University
Ames, IA 50011,
USA
tel: +515-294-7819
fax: +515-294-6814
e-mail: dandoug@iastate.edu

(4) Ari Huhta
Language Centre for Finnish Universities
University of Jyväskylä
P.O. Box 35
40351 Jyväskylä
Finland
tel: +358-41-603 539
fax: +358-41-603 521
e-mail: huhta@tukki.jyu.fi

(5) Carolyn Turner
Department of Second Language Education
McGill University
3700 McTavish Street
Montréal, Québec H3A 1Y2
Canada
tel: +514-398-6984
fax: +514-398-4679
e-mail: cx9x@musica.mcgill.ca

(6) Elaine Wylie
NLIIA, LTACC
Griffth University
Kessels Road
Nathan, Brisbane, Queensland 4111
Australia
tel: +617-875-7088
fax: +617-875-7090
e-mail: e.wylie@ais.gu.edu.au

APPENDIX THREE: The TFTS Textbase [TFTSTB]

==========

AU01: Australia: "Ethical Considerations in Ed. Testing..."

Record: AU01
TFTSmem: Wylie
Country: Australia
Auth/Pub: Masters, G. N., Australian Council for Educational Research Ltd (ACER).
Title(s): Ethical Considerations in Educational Testing. (1992) Issues Paper 2. (5 pages)
Language: English
Contact: Dr G Masters, Associate Director (Measurement) ACER, Private Bag 55, Camberwell, Victoria, Australia 3124. Fax: (613) 9277-5500
Govt/Priv: priv
Lang/Ed: ed
StanDef: guideline

Comments:

Objective:
In response to a request from the ACER Council, this paper outlines ethical considerations in the development and use of educational tests at ACER.

Intended audience/target group(s):

Developers of ACER tests and testing programs:
- external (to ACER) authors of tests published by ACER
- ACER staff developing tests for commercial sale, special-purpose tests for contracting clients, and ACER testing programs (e.g. scholarship testing programs).
- members of policy advisory groups established to oversee ACER test development.

Users of ACER tests and testing programs:
- teachers and schools who purchase and use ACER's commercial tests
- users of ACER's testing programs (e.g. scholarship testing, adult admissions testing)
- agencies or government departments commissioning special-purpose test development.

Procedure:
Based on Code of Fair Testing Practices in Education (1988) developed by a Joint Committee of the American Educational Research Association, the American Psychological Society, and the National Council on Measurement in Education.

Scope of influence:
Presumably endorsed by all developers and users.

Summary:
Relates to assessment in general. Outlines responsibilities for
Developers:
- developing/selecting appropriate tests
- interpreting test results
- ensuring fairness
- informing students
- seeking feedback

Users:
- selecting appropriate tests
- interpreting and reporting test results
- ensuring fairness
- informing students
- encouraging test takers to give feedback to ACER

Commentary:
This document has far-reaching applicability.

==========

AU02: Australia: "Assessment, referral and placement. No 17..."

Record: AU02
TFTSmem: Wylie
Country: Australia
Auth/Pub: Haughton, H. (ed). Commonwealth Department of Employment, Education and Training (DEET).
Title(s): Assessment, referral and placement. No. 17 in series Good Practice in Australian Adult Literacy and Basic Education (ALBE) (1992). (16 pages)
Language: English
Contact: Victorian Adult Literacy and Basic Education Council Inc. Fax: (613)-9654-1321.
Govt/Priv: govt
Lang/Ed: lang
StanDef: guideline

Comments:

Objective:
Aim of series is to promote good practice in ALBE. Funded under Australian Language and Literacy Policy (ALLP).

Intended audience/target group(s):
Teachers, teacher educators, program administrators.

Procedure:
There is an editorial committee drawn from ALBE experts in vari-

ous Australian states and the Australian Capital Territory. Editorial policy states that items chosen for publication describe activities which have wide acceptance amongst experienced practitioners.

Scope of influence:
As with most publications funded by DEET, there is a disclaimer stating that the opinions stated are not necessarily those of DEET. In this particular issue, the editor states that where contributors set out their credos explicitly, these are congruent with those enunciated by the Australian Council of Adult Literacy (ACAL) in their report on the 1991 Annual Conference and more recently in ACAL's position paper on the Australian Literacy and Numeracy (ALAN) scales.

Summary:
There are nine short articles on assessment, referral and placement in ALBE contexts (sometimes, but not always, explicitly ESL, since the ALLP tends not to differentiate at program delivery level). These articles are largely descriptive of approaches to assessment developed at state, industry or college level. The editor's introduction articulates an intention to seek guiding principles whereby "... student centredness (which) is one criterion by which good practice may be evaluated" can remain in focus when "... so many factors underlying current developments encourage practitioners to think in generalizable terms" and when "... current interest (in assessment) owes as much to political and economic influences as much as to educational ones".

==========

AU03: Australia: "School Certificate Grading System: Course..."

Record:	AU03
TFTSmem:	Wylie
Country:	Australia
Auth/Pub:	Board of Studies, New South Wales
Title(s):	School Certificate Grading System: Course Performance Descriptors. No date. 4-page leaflets for each of a number of languages, viz. Arabic, Chinese, Dutch, French, German, Greek (Classical), Greek (Modern), Hebrew, Indonesian, Italian, Japanese, Korean, Latin, Russian, Spanish, Turkish, and Vietnamese.
Language:	English
Contact:	Ms Hilary Dixon, Assessment Officer (LOTE), Board of Studies, PO Box 460, North Sydney, NSW, Australia 2059. Fax: (612) 9955 3557
Govt/Priv:	govt
Lang/Ed:	lang
StanDef:	performance
Comments:	

Objective:
To assist teachers in establishing an assessment program for the School Certificate (Year 10), viz. setting tasks (including formal tests) to elicit language behaviour from their students and relating student achievement to Performance Descriptors.

Intended audience/target group(s):
Teachers of languages other than English in NSW schools.

Procedure:
The booklets relate to the respective syllabus documents. In complementary documents for French and German, which provide exemplars of language behaviour at various levels of achievement, input from the respective syllabus committees is acknowledged.

Scope of influence:
The Board of Studies has the responsibility for the provision of a curriculum and assessment/examination of the curriculum for all students undertaking school education in the state of NSW from Years K (Kindergarten) to 12.

Summary:
Each leaflet outlines the types of tasks and the scheduling of tasks throughout Year 10 to allow students to demonstrate their maximum level of achievement. It provides descriptors at five levels for Listening, Speaking, Reading, Writing, and cultural aspects, and a guide for using these to award grades.

Commentary:
The documents are undated, but the descriptors are "For implementation in Year 10 in 1991", except Korean and Vietnamese, which are for 1992.

==========

AU04: Australia: "Subject Manual 5A: Languages other than ..."

Record: AU04
TFTSmem: Wylie
Country: Australia
Auth/Pub: Board of Studies, New South Wales
Title(s): Subject Manual 5A: Languages other than English (128 pages) and Subject Manual 5B: Languages other than English (117 pages). (1994).
Language: English
Contact: Ms Hilary Dixon, Assessment Officer (LOTE), Board of Studies, PO Box 460, North Sydney, NSW, Australia 2059. Fax: (612) 9955 3557
Govt/Priv: govt
Lang/Ed: lang
StanDef: guideline

Comments:

Objective:
To provide information pertaining to the Higher School Certificate Examination in 1995 and thereafter. (The Higher School Certificate is taken by students at the end of Year 12, which is the final year of formal school studies in NSW.)

Intended audience/target group(s):
School principals and teachers, examiners.

Procedure:
Changes to the rules in the Manuals are notified to schools through official notices in the Board Bulletin.

Scope of influence:
The Board of Studies has the responsibility for the provision of a curriculum and assessment/examination of the curriculum for all students undertaking school education in the state of NSW from Years K to 12.

Summary:
There are a Course Description, Assessment Guidelines (components and weightings, and suggestions for assessment tasks for the School Assessments component) and Examination Specifications for each of Chinese, Chinese for Students Educated through the Language (SETL), Classical Greek, French, German, Hebrew, Indonesian, Indonesian, Indonesian (SETL), Italian, Japanese, Japanese (SETL), Latin, Malay (SETL), Modern Greek, Spanish, and Vietnamese (5A) and Arabic, Armenian, Croatian, Czech, Dutch, Estonian, Hungarian, Khmer, Korean, Korean (SETL), Latvian, Lithuanian, Macedonian, Maltese, Polish, Russian, Serbian, Slovenian, Swedish, Thai, Turkish, Ukrainian (5B).

Commentary:
Manual 5B essentially covers languages which are offered under a national scheme which provides inter alia curriculum and assessment for less commonly taught languages, the National Assessment Framework for Languages at Senior Secondary Level (NAFLaSSL).

==========

AU05: Australia: "Educational objectives being tested in the..."

Record: AU05
TFTSmem: Wylie
Country: Australia
Auth/Pub: Australian Council for Educational Research Ltd. (ACER)
Title(s): Educational objectives being tested in the Commonwealth Secondary Scholarship Examination. 1967. (17 pages).
Language: English

Contact: Dr. Susan Zammit, ACER, Private Bag 55, Camberwell, Victoria, Australia 3124. Fax: (613) 9277-5599.
Govt/Priv: priv
Lang/Ed: ed
StanDef: test

Comments:

Objective:
To outline the background to the development of, and the rationale behind, the Commonwealth Secondary Scholarship Examination (CSSE) and to encourage informed consideration of the papers as models of examining.

Intended audience/target group(s):
Those who prepared and those who used the CSSE.

Procedure:
The document was prepared in 1966 and amended in 1977 by officers of ACER. The influence of Bloom's Taxonomy (1956) is acknowledged, as are statements prepared in connection with the IAEA Mathematics Project (International Evaluation of Educational Achievement, UNESCO Institute for Education, Hamburg).

Scope of influence:
The CSSE was used by all states of Australia as the basis for awarding scholarships to students in their final two years of secondary education. Some states used the exam from 1964 to 1974; some started to use it in 1965 or 1966.

Summary:
Outlines the background and rationale of the CSSE (and in particular its focus on intellectual ability rather than on the content of any particular prescribed syllabus). States the objectives for each of the four broad areas tested, viz Written Expression, Quantitative Thinking, Comprehension and Interpretation in the Sciences, and Comprehension and Interpretation in the Humanities. Further detail varies according to the area (e.g. for Written Expression the criteria for assessment are stated, for Quantitative Thinking the nature of the field covered is defined). There are informal specifications for tests for each of the four broad content areas. Notes that, because of practical limitations of large-scale examining, listening and oral communication (in English) can not be included, and that the exam can not allow students to demonstrate abilities in (inter alia) foreign languages.

==========

AU06: Australia: "Australian Scholastic Aptitude Test ..."

Record: AU06
TFTSmem: Wylie

Country: Australia
Auth/Pub: Australian Council for Educational Research Ltd. (ACER).
Title(s): Australian Scholastic Aptitude Test: Test Specifications. No date. (5 pages).
Language: English
Contact: Dr Susan Zammit, ACER, Private Bag 55, Camberwell, Victoria, Australia 3124. Fax: (613) 9277-5599
Govt/Priv: priv
Lang/Ed: ed
StanDef: guideline

Comments:

Objective:
To articulate the obligations of developers and users of the test.

Intended audience/target group(s):
Test developers (ACER officers) and users (see AU05).

Procedure:
Specifications prepared by the ACER in consultation with the users of the test (see 7) and ratified by the users.

Scope of influence:
The ASAT test was originally commissioned by the Commonwealth Government in 1970. In that year it was used by educational authorities in all Australian states except Victoria, and continues to be used regularly in a number of states and territories. After 1974 it was funded by the users.

Summary:
The document gives a general description of the test ("a 3-hour objective test of 100 questions") and indicates its purpose (to predict success in tertiary studies). It outlines the test content in terms of disciplines (drawn from humanities, social sciences, mathematics and sciences, but carefully avoiding the specific content of Year 11 and 12 syllabuses), appropriate types of stimulus material, and item content (the types and range of skills/abilities required overall and to answer any item). The construct of 'scholastic aptitude' is defined, and assumptions which underlie its validity are stated. Other aspects of validity and reliability are discussed. Finally, matters of copyright, security, informing prospective candidates about the test, reporting on performance of the test as a whole and of particular units and items, and test regeneration are covered.

Commentary:
The form of the document would suggest that it dates from relatively early in the history of the test.

==========

AU07: Australia: "NAFLaSSL Information Manual, 1993"

Record: AU07
TFTSmem: Wylie
Country: Australia
Auth/Pub: National Assessment Framework for Languages at Senior Secondary Level (NAFLaSSL)
Title(s): NAFLaSSL Information Manual. 1993. (115 pages).
Language: English
Contact: NAFLaSSL Coordinator, Board of Studies, PO Box 460, North Sydney, NSW, Australia 2059. Fax: (612) 9955 3557
Govt/Priv: govt
Lang/Ed: lang
StanDef: guideline

Comments:

Objective:
To provide information about the NAFLaSSL project, including instructions for the national management and administration of syllabuses and assessments.

Intended audience/target group(s):
Educational administrators and assessors.

Procedure:
The NAFLaSSL project was developed by a National Reference Group comprised of representatives of the assessment and accreditation authorities of all States and Territories of Australia, and assisted by representatives of State/Territory education departments, the Catholic Education Office, and the Independent Schools' Association. The project was funded by the Commonwealth Department of Employment, Education and Training (DEET) and consistently supported by the Australian Conference of Assessment and Certification authorities (ACACA).

Scope of influence:
All States and Territories participate.

Summary:
The document outlines the principles of 'nationalness', including planning and rationalisation to provide senior secondary school students in Australia with access to the widest possible range of languages and levels, and collaboration between all Australian assessment and curriculum authorities in the development, evaluation and redevelopment of syllabuses and associated assessment and reporting procedures. Subsequent sections address instructions for the administration of national syllabuses and assessments, guidelines for setters and vetters of national assessments, guidelines for presiding officers of national examinations, code of ethics, guidelines for reporting on national examinations, sample examination paper and mark sheets, grade

descriptions, and specimen assessments and support materials.

==========

AU08: Australia: "ACCESS Test Specifications..."

Record:	AU08
TFTSmem:	Wylie
Country:	Australia
Auth/Pub:	National Centre for English Language Teaching and Research (NCELTR), Macquarie University. Commonwealth Department of Immigration and Ethnic Affairs.
Title(s):	ACCESS Test Specifications. (Australian Assessment of Communicative English Skills) No date. Canberra, ACT: Department of Immigration and Ethnic Affairs. (37 pages).
Language:	English
Contact:	Professor D.E. Ingram, Director, Centre for Applied Linguistics and Languages, Griffith University, Nathan, Queensland, Australia 4111. Fax: (617) 3875 7090. e-mail: D.Ingram@ais.gu.edu.au
Govt/Priv:	govt
Lang/Ed:	lang
StanDef:	guideline
Comments:	

Objective:
To specify details of the assessment system.

Intended audience/target group(s):
ACCESS test developers and chief administrators (ACCESS is a secure test; the document is for restricted circulation).

Procedure:
The specifications were written in 1993 by a test development team managed by the NCELTR.

Scope of influence:
The ACCESS Test is designed specifically to meet the requirements of the Department of Immigration and Ethnic Affairs, and test development and administration are funded by the Department.

Summary:
The document describes and prescribes parameters for all aspects of the development and administration of the ACCESS test. There is a separate section for each of the (sub)tests of Listening, Oral Interaction, Reading and Writing, covering the nature of the (sub)test, its content, the proficiency levels against which it is referenced (and the range at which the majority of the items are pitched), and test structure.

==========

AU09: Australia: "Occupational English Test for Overseas..."

Record: AU09
TFTSmem: Wylie
Country: Australia
Auth/Pub: National Languages and Literacy Institute of Australia (NLLIA) Ltd.
Title(s): Occupational English Test for Overseas Qualified Health Professionals. No date. (2-page document for overseas audiences and a 3-page document for Australia audiences.)
Language: English
Contact: Ann Latchford, Project Officer OET, NLLIA Ltd Victorian Office, Level 9, 300 Flinders Street, Melbourne, Australia 3001. Fax: (613) 9629 4708.
Govt/Priv: priv
Lang/Ed: lang
StanDef: test

Comments:

Objective:
To inform about the background, purpose and nature of the OET, and the timing and locations of examination sessions.

Intended audience/target group(s):
Candidates and test users.

Procedure:
The document has been written by officers of the NLLIA.

Scope and influence:
The test described in the document is endorsed by most of the national and state Registration Boards/professional associations of doctors, dentists, nurses, physiotherapists, occupational therapists, speech pathologists, veterinarians, dietitians, radiographers, podiatrists, and pharmacists.

Summary:
The document describes briefly the background of the OET, viz. the transfer of its administration in 1991 from the National Office of Overseas Skills Recognition (NOOSR) within the Australian Department of Employment, Education and Training (DEET) and the needs analysis which led to the present form of the test. It states the purpose of the test, viz. to assess the English language proficiency of candidates seeking admittance to training programs and examinations, and outlines the content of the two non-profession-specific subtests (for reading and listening) and the two profession-specific subtests (for writing and speaking) and the general process of assessing each. The document for Australian audiences also includes dates for the test, policy on

re-sitting, and fees.

Commentary:
In 1986 Alderson, Candlin, Clapham, Martin and Weir reviewed the existing test used for health professionals, which was a test of general proficiency, and recommended the creation of a test which would assess the ability of candidates to communicate effectively in the workplace (ref. Alderson, C. et al. 1986. Language Proficiency Testing for Migrant Professionals: New Directions for the Occupational English Test. A report submitted to the Council on Overseas Professional Qualifications by the Testing and Evaluation Consultancy Unit for English Language Education, University of Lancaster, UK). The new speaking and writing subtests were initially developed and validated in 1987 and 1988 by McNamara, with the help of specialist informants. (ref. McNamara, T.F. 1990. Item Response Theory and the Validation of an ESP Test for Health Professionals. Language Testing, 7(1) pp52-76.)

==========

AU10: Australia: "NLLIA Japanese for Tourism and Hospitality..."

Record:	AU10
TFTSmem:	Wylie
Country:	Australia
Auth/Pub:	National Languages and Literacy Institute of Australia (NLLIA) Ltd.
Title(s):	NLLIA Japanese for Tourism and Hospitality Test: Notes for Test Administration Centres. No date. 3 pages.
Language:	English
Contact:	Dr. Joseph de Riva O'Phelan, NLLIA Ltd, Level 2, 6 Campion Street, Deakin, ACT, Australia 2600. Fax: (616) 281 3069. Fax in 1998 will be: (616) 9281-3069.
Govt/Priv:	priv
Lang/Ed:	lang
StanDef:	test
Comments:	

Objective:
To ensure correct administration of the test.

Intended audience/target group(s):
Test administrators.

Procedure:
The document was written by NLLIA officers.

Scope of influence:
The content of the document must be observed by all administra-

tors. The test itself is supported and recognised by the Inbound Tourism Association of Australia, Tourism Training Australia, Japanese Tour Wholesalers Committee and Australian Tourism Industry Association.

Summary:
The document outlines how to prepare the language laboratory facilities for the test, and how to administer it.

Commentary:
The Japanese for Tourism and Hospitality Test is part of a suite of tests. A test of Japanese for Tour Guides has been developed and a test of Korean for Tourism and Hospitality is being developed and trialled by officers of the NLLIA Language Testing Research Centre at the University of Melbourne.

==========

AU11: Australia: "Direct testing of general proficiency ..."

Record: AU11
TFTSmem: Wylie
Country: Australia
Auth/Pub: Wylie, E. and D.E. Ingram, National Languages and Literacy Institute of Australia (NLLIA) Ltd. Language Testing and Curriculum Centre, Griffith University.
Title(s): Direct testing of general proficiency according to the Australian Second Language Proficiency Ratings (ASLPR): Guidelines for the use of the ASLPR in testing. (1994). (10 pages).
Language: English
Contact: Elaine Wylie, Deputy Director, NLLIA Language Testing and Curriculum Centre, Griffith University, Nathan, Queensland, Australia 4111. Fax: (617) 3875 7090. e-mail: E.Wylie@ais.gu.edu.au
Govt/Priv: priv
Lang/Ed: lang
StanDef: guideline

Comments:

Objective:
To establish procedures for direct, adaptive rating of candidates on the ASLPR.

Intended audience/target group(s):
ASLPR testers

Procedure:
The guidelines have been prepared by the developers of the ASLPR to meet the needs of trainees and as a result of observation of trainees in practicum sessions.

Scope and influence:
The ASLPR is used by the Australian Departments of Immigration and Ethnic Affairs (DIAEA) and Employment, Education and Training (DEET), and a number of universities, professional registration bodies and
language teaching institutions throughout Australia.

Summary:
The document outlines issues of validity in relation to ASLPR testing. It describes overall procedures for selecting, preparing and administering tasks for learners at different levels of proficiency and different backgrounds, with separate sections for productive and receptive language behaviours. The final section addresses issues of reliability and practicality (e.g. level of resource allocation as a function of the stakes involved in a particular test administration).

Commentary:
The document complements the ASLPR scale, a formal paper "Introduction to the ASLPR" by D. E. Ingram, and a set of videos, to provide a manual for developing and administering direct adaptive test tasks and rating behaviour elicited.

==========

AU12: Australia: "Discussion Papers 1-21 (1986/7/8)..."

Record: AU12
TFTSmem: Wylie
Country: Australia
Auth/Pub: Sadler, R. et al, Assessment Unit, Board of Secondary School Studies, Queensland.
Title(s): Discussion Papers 1-21 (1986/7/8) (details under Comments below)
Language: English
Contact: Ms A. Vitale, Board of Senior Secondary School Studies, P.O. Box 307, Spring Hill, QLD 4000. Fax: (617) 3832 1329.
Govt/Priv: govt
Lang/Ed: ed
StanDef: guideline

Comments:

Objective/Purpose:
To facilitate "the formulation of curriculum and assessment policy within secondary schools".

Target group/Audience intended:
The teaching profession in the state of Queensland.

Procedure:

The suite of papers was written by members of the Assessment Unit of the BSSS under the leadership first of Royce Sadler and subsequently Warren Beasley. The papers address concerns expressed by teachers during the early years of criteria-based assessment in Queensland (later called 'standards-based assessment' or 'criteria- and standards-based assessment'). Each paper carries a note inviting teachers to react to and comment on its contents.

Scope of influence:
The Board of Secondary School Studies (BSSS) (now the Board of Senior Secondary School Studies) has responsibility for curriculum and assessment for Years 11 and 12, the last two years of formal schooling in Queensland. The papers have played an invaluable role in speaking to teachers in "a bold and imaginative ground-breaking exercise" of introducing criteria- and standards-based assessment in Queensland schools (Campbell, W. J. et al. 1983. Implementation of ROSBA in Queensland Secondary Schools. Department of Education, University of Queensland, p19.) See also Commentary.

Summary:
The Abstract of each paper is reproduced below after the publication details.

Discussion Paper 1: Sadler, R. 1986. ROSBA's Family Connections. (8 pages). The Radford scheme belonged to a family of procedures known technically as 'norm-referenced' assessment. The current system, called ROSBA, focuses on criteria and standards and belongs to the 'criterion-referenced' family. In this Paper, something of the similarities and differences between these two families are outlined. It is also shown how ROSBA differs from the criterion-referenced testing movement in the U.S.A.

Discussion Paper 2: Sadler, R. 1986. The Case for Explicitly Stated Standards. (6 pages). Nine reasons for making criteria and standards explicit are outlined in this paper. The first six set out general benefits of being specific; the final three make a case for having explicit statements incorporated into syllabus documents.

Discussion Paper 3: McMeniman, M. 1986. A Standards Schema. (7 pages). This paper presents a model for pegging standards along different criteria or performance dimensions relevant to a particular subject area. The model stands in contrast with mastery-learning forms of criterion-referenced assessment. It shows also how standards can be combined for the award of exit Levels of Achievement.

Discussion Paper 4: Sadler, R. 1986. Defining Achievement Levels. (10 pages). Statements of the five achievement levels (VLA, LA, SA, HA, VHA) constitute an important element of school Work Programs under ROSBA. In this Paper, some of the things to avoid in writing good achievement level state-

ments are outlined. The treatment is necessarily general, but the broad principles are applicable to all subjects in the curriculum.

Discussion Paper 5: Sadler, R. 1986. Subjectivity, Objectivity, and Teachers' Qualitative Judgments. (10 pages). Qualitative judgments play an essential role in the assessment of student achievements in all subjects. This Discussion Paper contains a definition of qualitative judgments, a discussion of the meanings of subjectivity and objectivity, and a statement of certain conditions that must be satisfied if qualitative judgments are to enjoy a high level of credibility.

Discussion Paper 6: McMeniman, M. 1986. Formative and Summative Assessment - A Complementary Approach. (7 pages). This paper attempts to describe where formative and summative assessment might most efficiently be applied under ROSBA. It touches also on one of the major concerns of ROSBA that summative assessment of students should not rely solely or even principally on one-shot examinations. In support of this concern, a rationale is presented for a series of student performances being used as the basis for judging whether summative assessments accurately reflect the real capabilities of students.

Discussion Paper 7: Findlay, J. 1986. Mathematics Criteria for Awarding Exit Levels of Achievement. (17 pages). This paper is about the criteria for judging student performance in mathematics at exit. The particular mathematics course considered is the current Senior Mathematics, although much of the paper has relevance for Mathematics in Society, and for Junior Mathematics. The first part of the paper considers some problems associated with current assessment practices, in terms of the syllabus and school translations of it. Then some contributions from different sources to the search for criteria are discussed. In the third section criteria are defined, and along with standards, organised as a possible model for awarding exit Levels of Achievement. The final section of the paper contains a discussion of the model and some possible implications. The model itself is included as a separate appendix.

Discussion Paper 8: Sadler, R. 1986. Developing an Assessment Policy Within a School. (11 pages). This Discussion Paper contains a number of guidelines of potential interest to teachers and school administrators who wish to develop an internal school policy on assessment. The five principles outlined relate to the quantity and type of information required, the structure of subjects in the curriculum, and the needs of students and parents for feedback. The Paper concludes with suggestions for two quite different ways of responding to a desire for reform.

Discussion Paper 9: Sadler, R. 1986. General Principles for

Organising Criteria. (9 pages). Criteria are, by definition, fundamental to a criteria-based assessment system. This Paper outlines a conceptual framework for organising criteria, using a particular subject as an example. It also shows how student achievement can be recorded using criteria.

Discussion Paper 10: Sadler, R. 1986. Affective Objectives Under ROSBA. (9 pages). Affective objectives have to do with interests, attitudes and values, and constitute an important aspect of education. They have implications for teaching, learning, the curriculum, and the organisation of schooling. Whether they should be assessed, either at all or in specified areas, is an issue worthy of some discussion. In this Paper, it is argued that affective responses should not be incorporated into assessments of achievement.

Discussion Paper 11: Sadler, R. 1986. School-based Assessment and School Autonomy. (8 pages). School-based assessment in Queensland means that teachers have responsibility for constructing and administering assessment instruments and for appraising student work. But because certificates are issued from a central authority, the assessments must be comparable from school to school. In addition to being school-based, the ROSBA system is criteria-based as well. It is argued in the Paper that using uniform criteria and standards across the state allows for variety of approach in assessment and helps to achieve comparability without destroying the autonomy of the school.

Discussion Paper 12: Sadler, R. 1987. Defining and Achieving Comparability of Assessments. (11 pages). Four interpretations of comparability are identified and discussed in this Paper: comparability among subjects, students, classes, and schools. The last of these, comparability among schools in each subject is a crucial concern in a school-based assessment system in which certificates are issued by a central authority. Only when the Levels of Achievement have consistent meaning across the state can public confidence in the certificate be maintained. It is argued here that achievement of comparability is fully compatible with the concept of teachers as professionals, and with accountability of the profession to the public at large.

Discussion Paper 13: McMeniman, M. 1987. Towards a Working Model for Criteria and Standards under ROSBA. (10 pages). This paper is concerned with clarifying the use under ROSBA of the terms 'criterion' and `criteria' and with arguing the case for specifying standards within this nomenclature. A model is then presented of how criteria and standards might ideally operate under ROSBA.

Discussion Paper 14: Bingham, R. 1987. Criteria and Standards in Senior Health and Physical Education. (15 pages). This paper is concerned with the assessment of students' global

achievements in Senior Health and Physical Education. It examines the notion of global achievement in the subject and suggests the criteria and standards by which the equality of student performance can be judged. It also suggests specifications for awarding exit Levels of Achievement which reference the standards schema.

Discussion Paper 15: Findlay, J. 1987. Improving the Quality of Student Performance through Assessment. (8 pages). Two basic assessment mechanisms through which the quality of student performances can be improved are feedback and information supplied about task expectations prior to performance. In this paper the complementary nature of these two mechanisms is examined, while feedback is analysed to indicate the value of certain forms of feedback over others. The paper complements and further develops some ideas concerned with formative and summative assessment presented in an earlier Discussion Paper.

Discussion Paper 16: Beasley, W. 1987. A Pathway of Teacher Judgments: From Syllabus to Level of Achievement. (9 pages). This paper traces the decision-making process of teachers which allows information about student achievement within a course of study to be profiled over time. It attempts to place in perspective the different levels of decision-making and identifies the accountability of such judgments within the accreditation and certification process.

Discussion Paper 17: Beasley, W. 1987. Assessment of Laboratory Performance in Science Classrooms. (9 pages). Laboratory activity in high school science classrooms serves a variety of purposes including psychomotor skill development and concept introduction and amplification. However it does not by itself provide sufficient experience that the learning outcomes are direct in the case of concept development, or of a high order in the case of psychomotor skill. This paper sets out a schema of global outcomes which might provide a more realistic framework for decisions about students laboratory performance at the end of a course of study. The use of the schema within the total course of study to award an exit Level of Achievement is also discussed.

Discussion Paper 18: Beasley, W. 1987. Profiling Student Achievement. (9 pages). The record of student performance over a course of study provides a summary of information from which ultimately a judgment is made by a teacher on an appropriate exit Level of Achievement. This summary, determined from and including qualitative and/or quantitative statements of performance, captures sufficient information to indicate standards of achievement. The judgments of standards attained are reference initially to interim criteria at the end of semester, and finally to global criteria at the completion of a course of study. The design characteristics of a format to profile records of student performance are discussed.

Discussion Paper 19: Findlay, J. 1987. Principles for Determining Exit Assessment. (9 pages). Six key principles underpin exit assessment. They are continuous assessment, balance, mandatory aspects of the syllabus and significant aspects of the course of study, selective updating, and fullest and late information. These principles are explained in some detail and related to courses of study.

Discussion Paper 20: Findlay, J. 1987. Issues in Reporting Assessment. (7 pages). This paper is about some of the issues associated with reporting in schools. Provided are a number of suggestions which may form a partial solutions to problems which arise.

Discussion Paper 21: Sadler, R. 1988. The Place of Numerical Marks in Criteria-Based Assessment. (10 pages). Numerical marks form the currency for almost all assessments of student achievement in schools, and the use of them is rarely if ever challenged. In this paper, a number of assumptions underlying the use of marks are identified, and the appropriateness of marks in criteria-based assessment is examined. The conclusion is drawn that continued use of marks is more likely to hinder than to facilitate the practice of judging student achievements against fixed criteria and standards.

Commentary:
Queensland was the first Australian state to abolish external examinations, and from the early 1970s until 1981 assessment was school-based but norm-referenced. Criteria-based assessment was phased in between 1981 and 1987 (ref. Connell, W.F. 1993. Reshaping Australian Education. Hawthorn, Victoria: The Australian Council for Educational Research.)

The papers are written in a very accessible style. A number of them, particularly those by Sadler, have been widely cited in the literature on educational assessment in other Australian states.

==========

AU13: Australia: "Tasmanian Certificate of Education..."

Record:	AU13
TFTSmem:	Wylie
Country:	Australia
Auth/Pub:	The Schools Board of Tasmania.
Title(s):	Tasmanian Certificate of Education (1994) (173 pages)
Language:	English
Contact:	Graham Fish, Chief Executive Officer, P.O. Box 147, Sandy Bay, Tasmania 7006. Fax: (610) 224 9175. After Nov. 1996, fax will be (613) 6224 9175
Govt/Priv:	govt

Lang/Ed: ed
StanDef: guideline

Comments:

Objective/Purpose:
To introduce the Tasmanian Certificate of Education.

Target group/Audience intended:
Teachers, educational administrators.

Procedure:
Not stated

Scope of influence:
The Schools Board of Tasmania oversees the process of certification in the senior secondary years.

Summary:
After a short background statement, which includes a comment about the combined external/internal system operating in Tasmania, there is a 16-page section on Assessment and Moderation. The sub-section "Guidelines for Internal Assessment" indicates that a variety of methods may be used for internal assessment, including observation and tests, and notes that specific tasks need not be devised for each criterion stated in the syllabus. There are also statements about recording and reporting internal assessments. A further section lists the syllabuses for each subject. Languages available are Chinese, English, English as a Second Language, French, German, Indonesian, Italian, Japanese, Latin, Russian, Spanish.

Commentary:
The accompanying letter states that syllabuses (which were not sent) incorporate standards documents and exemplars, with the most extensive support material available for the language syllabuses in English, ESL, French, German, Indonesian, Italian and Japanese.

==========

AU14: Australia: [Three documents from the Sec. Ed. Authority]

Record: AU14
TFTSmem: Wylie
Country: Australia
Auth/Pub: Secondary Education Authority (SEA), Western Australia.
Title(s): (i) Assessment, Grading and Moderation Manual - 1994 (22 pages)
(ii) Tertiary Entrance Examinations Examiners' Handbook. (1994) (17 pages plus appendices)
(iii) Curriculum Area Framework: Languages other

than English. (undated) (33 pages)
Language: English
Contact: Dr Bob Peck, A/Senior Education Officer (Assessment) Secondary Education Authority, 27 Walters Drive, Herdsman Business Park, Osborne Park, Western Australia 6017. Fax: (619) 273 6301. After Sep., 1997, fax will be (619) 9273 6301.
Govt/Priv: govt
Lang/Ed: ed
StanDef: guideline

Comments:

Objective/Purpose:
(i) To inform about SEA policy and procedures relating to assessment, grading and moderation in Year 11 and Year 12 SEA Accredited Courses.
(ii) To describe and prescribe the role of examiners for public examinations in approved Tertiary Entrance Score (TES) subjects
(iii) (inter alia) To assist in planning for course development and evaluation at the upper secondary level, so as to cater more effectively for the needs of all students. (One of the criteria for evaluation is the appropriateness of objectives, content and assessment to Year 11 or Year 12 level)

Target group/Audience intended:
(i) Upper secondary school teachers and school administrators
(ii) Examiners
(iii) "The primary audiences for the Curriculum Area Frameworks are SEA Committees such as the Curriculum Area Committees and the Syllabus Committees. However, other individuals involved in course development at Year 11 and Year 12 level will also use these documents in preparing courses for accreditation." (p2)

Procedure:
(i) Document prepared by SEA officers. The section 'Developing a School Assessment Policy' was reviewed in consultation with Principals during 1993.
(ii) Document prepared by SEA officers.
(iii) The Frameworks were prepared during 1992 by a writer or team of writers with extensive experience and awareness of local and national trends in the curriculum area. The writers worked within parameters established by the SEA and reported to the SEA Curriculum Area Committees which are comprised of representatives of education providers, post-secondary institutions, industry and community groups. The Curriculum Area Committees work closely with the Syllabus Committees, with the former having a management coordination focus and the latter having a subject expertise focus.

Consultation drafts of Frameworks were distributed to schools, colleges and post-secondary institutions for comment. Feedback received was collated and considered by the Curriculum Area Committees in finalising each Framework for SEA endorsement.

Scope of influence:
(i) SEA has statutory responsibility for ensuring comparability of grades and numerical assessments within and between schools. Schools offering SEA Accredited Courses must be familiar with the policy and procedures outlined in the Manual.
(ii) SEA conducts public examinations in approved Tertiary Entrance Score (TES) subjects and provides information to tertiary institutions regarding the performances of students seeking entry to those institutions.
(iii) as for (i) and (ii)

Summary:
Documents (i) and (ii) refer to assessment in general, not specifically languages. Document (iii) addresses languages other than English.

(i) The document has six sections, Assessment and Grading Requirements; Developing a School Assessment Policy; Assessment and Grading in Accredited Courses; Assessment of Students with Disabilities; Inclusion of Assessment in Out-of-school Learning Situations for Grading in Accredited Courses; and Comparability and Moderation. Appendices give details of requirements for a 'Moderation Visit' by an SEA officer to a school, e.g. providing copies of the accredited school assessment program, all assessment instruments with keys if appropriate, and samples of assessed student work to illustrate cut-off points for grades.

(ii) The document features a Code of Conduct for TEE Examiners and independent reviewers. This includes references to declaring any potential conflict of interest, to maintaining confidentiality, and to maintaining a low public profile, and general guidelines on fairness and impartiality. A major section of the document addresses the setting of papers, and includes mechanical and managerial aspects, e.g. formatting and timing, as well as aspects of validity (e.g. covering the range of the syllabus and reflecting the objectives of the course; eliminating bias due to focus on mental processes or knowledge limited to particular subgroups of candidates; and wording rubrics unambiguously).

(iii) One of the appendices has a section on assessment, defined as "the ongoing process of collecting information about student achievement and performance and making decisions based on that information" (p23). It stresses that assessment is integral to course development and the learning process. It states that assessment tasks, criteria for judging performance and marking procedures should give "primary and consistent significance to the purposeful use of language" (p24). Having stressed that each course of study has its own set of assessment structures, which must be adhered to by teachers, it outlines parameters for emphases in assessment on different content areas and different learning outcomes. It specifies that assessment may include examinations and/or major tests and tasks based on the purposeful use of languages, listing task-types such role play, problem-

solving tasks, and information gap activities as appropriate types of assessment.

Commentary:
Complementing the documents reviewed there are eleven Syllabus Manual Volumes, of which Volume III, Languages other than English (156 pages) was provided.

==========

AU15: Australia: "The Process of Assessment, Grading, and ..."

Record:	AU15
TFTSmem:	Wylie
Country:	Australia
Auth/Pub:	Griffith University
Title(s):	The Processes of Assessment, Grading and Dissemination of Results. Griffith University Calendar. Section 8.20: (1991 revision) (8 pages)
Language:	English
Contact:	Dr. Lyn Holman, Academic Registrar, Griffith University, Nathan, Queensland, Australia, 4111.
Govt/Priv:	priv
Lang/Ed:	ed
StanDef:	guideline

Comments:

Objective/Purpose:
To identify
- "the respective ambits of responsibility in assessment processes of the Academic Committee, the Education Committee, the Divisional Assessment Boards and sub-Boards, and faculty staff as part of course design and teaching teams and as examiners, supervisors and invigilators;
- "the composition, functions, and methods of operation of Divisional Assessment Boards and sub-Boards;
- "the policy on, and procedure for, appeals against Divisional Assessment Board procedures or determinations;
- "the nature of information on student performance which should be made available, the circumstances under which some or all of the information is released, and the machinery for its dissemination.

Target group/Audience intended:
All academic staff.

Procedure:
Not stated.

Scope of influence:
Applies to all internal assessment processes of Griffith University.

Summary:
There is an introductory statement that student academic assessment takes place within a framework of policy and procedures which is established by each Division within the framework of general University policy. Following sections outline the responsibilities and constitution of the various committees and (sub)boards, and the responsibilities of teaching teams, individual examiners, supervisors and invigilators. There are sections on appeals, marking and grading, dissemination of information on assessment results and disposal of assessment material.

Commentary:
Since 1991 the University has changed from Divisions to Faculties. The Academic Registrar in a personal communication to the TFTS member indicated that the University is currently revising all its assessment policies.

==========

AU16: Australia: "Faculty Assessment Policies" [Griffith Univ.]

Record:	AU16
TFTSmem:	Wylie
Country:	Australia
Auth/Pub:	Griffith University
Title(s):	Faculty Assessment Policies (various documents; see Comments below for detail)
Language:	English
Contact:	Dr. Lyn Holman, Academic Registrar, Griffith University, Nathan, Queensland, Australia, 4111.
Govt/Priv:	priv
Lang/Ed:	ed
StanDef:	guideline

Comments:

Objective/Purpose:
The Faculty of Commerce and Administration document states that their document has been prepared to assist faculty staff, particularly new staff, to gain a knowledge of the Faculty's policies with regard to assessment of undergraduate courses. The other documents do not have a purpose overtly stated.

Target group/Audience intended:
In all cases Faculty staff. In some cases students as well. The Faculty of Commerce and Administration specifically states that the document is not intended for issue to students; the Faculty of Education specifically states that the Assessment Policy (presumably a copy of the document) will be provided to each student on commencement of a program.

Procedure:

The Faculty of Humanities document states that the procedures in the document have been approved by the Faculty Standing Committee and can be varied only with that Committee's approval.

Scope of influence:
As stated in the Griffith University's Processes of Assessment, Grading and Dissemination of Results (record AU16) student academic assessment takes place within a framework of policy and procedures which is established by each Division (Faculty) within the framework of general University policy.

Summary:
Separate documents representing a number of Faculties were provided. These were the Faculty of Commerce and Administration Undergraduate Academic Board Assessment Policy (1993) (20 pages plus 2-page Attachment); the Division of Education Assessment Policy (1991) (6 pages plus 4-page Attachment); the Faculty of Humanities Assessment Practices and Policies (1993) (6 pages plus 1-page Attachment plus 4 pages of Guidelines); the Division of Nursing and Health Sciences Assessment Policies and Procedures (Gold Coast Campus)(1991) (6 pages). All documents refers to assessment in general, not specifically languages. The most comprehensive document, from Commerce and Administration, has separate sections on Assignments, Assessment of Joint Projects, Double Marking of Assessment Items, Examinations, Supplementary Assessment, Special Assessment, the Grade of PC (pass Conceded), Failure of Courses, Student Appeals Procedures, Procedure to Approve Assessment Changes, Assessment and the Award of Marks and Grades, Cheating, including Plagiarism, Recording of Marks and Timing of Assessment Board Returns, and Assessment of Post-Graduate Students in Undergraduate Courses. The appendix is on Group Assessment.

The Education Faculty document begins with a statement of Principles of Assessment. This states that the Faculty endorses criteria-based principles of assessment and considers that assessment has a role not only in certification but also as an aid to learning. Assessment practices for each course should reflect a diversity of assessment strategies to ensure the fairest and most comprehensive judgment of learning outcomes. It also states that assessment information and outcomes should provide information to assist in the continuing and periodic evaluation of program and course objectives, content, teaching methods and procedures.

Commentary:
Since 1991 the University has changed from Divisions to Faculties. The Academic Registrar in a personal communication to the TFTS member indicated that the University is currently revising all its assessment policies.

==========

AU17: Australia: "ESL Development: Language and Literacy in..."

Record:	AU17
TFTSmem:	Wylie
Country:	Australia
Auth/Pub:	McKay, P. et al. (see details under Summary, below) National Languages and Literacy Institute of Australia (NLLIA)
Title(s):	ESL Development: Language and Literacy in Schools Project (second edition 1994). Volume I: Teachers' Manual - Project Overview, NLLIA ESL Bandscales and Materials (273 pages). Volume II: Documents on Bandscale Development and Language Acquisition (259 pages).
Language:	English
Contact:	Dr. Penny McKay. School of Language and Literacy Education, Faculty of Education, Queensland University of Technology, Kelvin Grove Campus, Locked Mail Bag No. 2. Red Hill, Queensland, Australia 4059. Fax: (617) 864 3988. The volumes can be purchased from NLLIA Ltd Victorian Office, Level 9, Flinders Street, Melbourne, Australia 3001. Fax (613) 9629 4708.
Govt/Priv:	priv
Lang/Ed:	lang
StanDef:	performance

Comments:

Objective/Purpose:
Volume I contains the major outcomes of the project (see Summary below), which are intended to provide teachers with a reference and guidelines for assessing and reporting the development of ESL learners in schools. Volume II contains documents which provide the applied linguistic and educational background to the project.

Target group/Audience intended:
Teachers, teacher educators, program planners and policy-makers.

Procedure:
The project was managed by the NLLIA Directorate (Director, Joseph Lo Bianco) and undertaken by three research and development centres of the NLLIA, the NLLIA Language Testing and Curriculum Centre at Griffith University, the (then) Language Testing Centre at Melbourne University, and the NLLIA Language Acquisition Research Centre at Sydney University. It was originally co-ordinated nationally by Patrick Griffin, who continued as a consultant to the project, and subsequently by Penny McKay. There was a strong consultative structure for the project, consisting of a Project Co-ordination Unit (relevant NLLIA personnel) a Steering Committee (members of the Project Co-ordination Unit, representatives of the Department of Employment, Education and Training and an independent member), and a Reference Commit-

tee (members of the Steering Committee, representatives of all state education departments/ministries, of the National Catholic Education Commission, of the National Council of Independent Schools Associations, of the Australian Education Council (AEC), the Australian Council for Educational Research, and the National Centre for English Language Teaching and Research. The Steering Committee and Reference Committee met four times during the project.

Scope of influence:
The Bandscales are used in a number of states for diagnostic, research and other purposes. It is hoped that assessments made on the Bandscales will be used as a means of allocating and distributing funds for ESL programs throughout Australia.

Summary:
Volume I has a number of sections which provide an overview of the project (e.g. terms of reference, project management, relationship of the various components).

There is a section called "Principles Informing the ESL Development Project", which stated the following six principles (each of which is elaborated in the actual document).

The ESL Development Project should

i) address the requirements of the political context
ii) draw on a broad philosophical and research base
iii) represent all learners as far as possible, and do so positively and equitably
iv) accommodate the realities of and the shared perspectives about ESL teaching and learning in the school ESL field
v) recognise the practical constraints of ESL teaching and support in schools
vi) emphasise the need for further research and development in school ESL.

There is a section "Using the ESL Development Project Materials", which provides brief guidelines for teachers in using the materials to assess and report (e.g. devising activities to obtain the most valid sample of language behaviour, systematic recording of observations). There is another section "The Development of the ESL Bandscales", which inter alia discusses the theoretical models underpinning the Bandscales, addresses construct, content and face validity, and outlines the limitations of the Bandscales and the need for further research.

Major sections contain the set of Bandscales for Junior Primary Middle/Upper Primary, and Secondary phases of schooling (McKay, Sapuppo and Hudson), and exemplar assessment activities for all phases and observation guides in the four macroskills to guide teachers towards effective in-class assessment (Lumley, Mincham and Raso) and reporting formats (Greco, Raso, Lumley and McKay).

Volume II contains documents which were significant in the developmental process, including the report from the NLLIA Language Acquisition Research Centre "An Empirical Study of Children's ESL Development and Rapid Profile" (Pienemann and Mackey).

==========

AU18: Australia: "Ethical Guidelines..." [Migrant Ed.]

Record:	AU18
TFTSmem:	Wylie
Country:	Australia
Auth/Pub:	Adult Migrant Education Services, Victoria
Title(s):	Ethical Guidelines. 1992. 6 pages.
Language:	English
Contact:	Mr Chris Corbel, AMES, 1st. Floor, Myer House, 250 Elizabeth Street, Melbourne, Victoria, Australia 3000. Fax: (613) 9663 1130.
Govt/Priv:	govt
Lang/Ed:	lang
Stan/def:	guideline

Comments:

Objective/Purpose:
"... to establish ethical guidelines for individuals and organisations involved in the provision of English language assessments. The primary concern of these guidelines is to ensure that the aims, functions and operations of language assessment services will enhance opportunities for individuals rather than to limit them."

Target group/Audience intended:
Individuals and organisations involved in the provision of English language assessments conducted by AMES Victoria.

Procedure:
The statement was developed by AMES Languages Assessment Services in consultation with the Office of the Equal Opportunity Commissioner, the Victorian Ethnic Affairs Commission, Ethnic Liaison Officers of the Australian Council of Trade Unions and the Victorian Trades Hall Council, the Division of Further Education at the Victorian Ministry of Education, the ESL Liaison Officer of the State Training Board and the ESL Department of the Footscray Campus of the Western Metropolitan College of TAFE (Technical and Further Education).

Scope of influence:
The guidelines apply to all parties to language assessments conducted by AMES Victoria

Summary:

Section 1 outlines the context of the document, viz. the fact that English is the language of the major and powerful institutions of Australian society, and the Australian Language and Literacy Policy (1991), which has as its primary goal the development and maintenance of English for all Australian residents and the establishment of education programs to meet their diverse learning needs.

Section 2 relates English language assessment to work training. It stipulates that assessments, whether for employed or unemployed persons, should not be carried out unless language audits are first and foremost the identification of training needs to those purposes has been clearly expressed and understood by participants in the individual assessment or audit, whose informed consent is necessary.

Section 3 outlines the need for ethical guidelines (see Objective/Purpose above)

Section 4 describes the two broad areas for which language assessment services are provided (individual assessments for government sponsored programs and audits for industry) and Section 5 outlines the various purposes of assessments within these areas.

Section 6 addresses implementation of language assessment services. It stipulates the formation of tripartite consultative committees (with management, union and program-provider representation) for assessments related to English language provision for industry and government departments. It addresses contractual agreements, and reporting formats (stipulating inter alia that reports must not include any information which the assessor is not qualified to give, e.g. statements on medical condition).

Section 7 covers the standards of qualifications and training of assessors, and the requirements that they participate in moderation processes and accept the ethical guidelines.

Section 8 covers the choice of assessments instruments and the shelf-life of assessments (six months). It stipulates that instruments chosen by assessors must satisfy certain requirements in terms of validity, reliability, and match with the purposes of the assessments. It further stipulates that individuals and organisations must know in advance the purposes for which the instrument was designed and the limitations of its use, and that they must have evidence of its validity. Notes to this section state that the Australian Second Language Proficiency Ratings (ASLPR) scale is extensively used by AMES Victoria, and outline how the ASLPR (and the related Exrater) meet the requirements of validity, reliability and match with the purposes of AMES assessments.

Commentary:
This document is of particular interest because of its focus on language audits for companies and industries. (Victorian AMES has conducted a number of such audits, including a major one for the automobile industry.) It provides a useful model for the implementation of such audits.

==========

Explanatory Notes regarding records from Canada

Education:
Canada is a bilingual country. All provinces (the English-speaking ones) except Québec have French as a Second Language (FSL) as the official second language. In Québec (the French-speaking province) ESL is the official second language.

Public Schcols:
Ministries of Education: In Canada, public school educational standards and standards of practice in testing (if they exist) are set at the provincial level by the Provincial Ministry of Education. In general the provincial exams (if they exist) reflect the curriculum and are developed by provincial educators, teachers, and experts. Manitoba is presently reintroducing provincial exams. They are presently working on a document for evaluation policies and procedures concerning English Language Arts and French-Mother Tongue curriculum. The intended audience includes all those involved in such language curriculum in the province. Documents and workshops are provided to teachers to help
them with classroom assessment.

From the English-speaking provinces, materials pertaining to "standards of performance" for ESL students being integrated into the school were received, but not materials pertaining to "standards of evaluation practices". For example, in Ontario, there is a draft version of a document entitled Provincial Standards: Language (Grades 3, 6, and 9). This comes from the Ministry of Education and Training Curriculum and Assessment Team and is part of the Provincial Standards Project.

The English-speaking provinces heard from mainly focused on the specific FSL tests used in their school systems, rather than on standards or guidelines. These provinces included British Columbia, Manitoba, New Brunswick, Newfoundland, Nova Scotia, Prince Edward Island, and Ontario. In addition, Québec referred to this List. Many of them said that their tests had been reviewed and published in the "Annotated List of French Tests: 1991 Update" in the Canadian Modern Language Review (CMLR).

Others said that they referred to this List when needing to select or construct a French test. So as not to be repetitive, the CMLR list is cited and summarized in CA03 below.

(Note: with the permission of CMLR this list was reproduced in Language Testing Update, issues 12 and 13.)

Besides Québec, only one province sent materials specific to ESL pertinent to our TFTS interest. This was in the form of curriculum guidelines with a section on assessment (Record CA04).

Evaluation in second language courses, French/English as a Second Language (FSL/ESL), is a particular preoccupation in the province of Québec, where two school systems exist side by side throughout the province, the French-speaking system and the English-speaking system. Language instruction of the respective second language is started in elementary school. The amount of Québec government documentation dealing with language evaluation is abundant. In addition, teachers and educators are provided with generic documents concerning evaluation. The Ministère de l'Éducation du Québec (MEQ) has provided a sampling of the major materials concerning guidelines and suggestions for testing and evaluation practice in Québec.

(Other provinces deal in general evaluation, but guidelines for evaluation specific to languages is rare.)

University and Adult Education:
Record CA17 is a published document. Several university/adult language programs said they had suggestions, but no formal guidelines or standards to assist their educators in test development and/or selection. Many of them did have in-house tests, however, but chose not to send them.

Regarding Canadian Universities, some consistency was found among professors and researchers in terms of what they use or rely upon as guidelines for the construction or development of language-related tests/assessment. Sometimes professors are contracted or asked to be consultants for a specific test development project by the government, Ministry of Education, or some institution. Often these institutions rely on the professors' expertise and as one professor mentioned, "...we get zero guidelines or documents to guide us". It is at this point that professors seem to have developed a list of references to help guide them. Below, is a list of the references that were repeatedly received. Repeatedly means at least four sources mentioned the reference (i.e., 4 professors from 4 different universities). These references are not all necessarily Canadian, they are what some Canadian professors are referring to:

Frederiksen, N., Mislevy, R.J., & Bejar, I.I. (Eds.). (1993). Test Theory for a New Generation of Tests. Hillsdale, NJ: Lawrence Erlbaum Associates.

Linn, R.L., Baker, E.L., & Dunbar, S.B. (1991). "Complex, performance-based assessment: Expectations and validation criteria". Educational Researcher, 20(8), 15-21.

Royer, J.M., Cisero, C.S. & Carol M.S. (1993). "Techniques and procedures for assessing cognitive skills". Review of Educational Research, 63(2), 201-243.

Bowd, A., McDougall, D. & Yewchuk, C. (1994). Educational Psychology for Canadian Teachers. Toronto, ONT: Harcourt Brace Canada. (The section I was referred to: "Part Four - Measurement and Evaluation of Achievement".)

Walberg, H.J. & Gaertel, G.D. (Eds.). (1990). The International Encyclopedia of Educational Evaluation. Oxford: Pergamon.

Language Proficiency Testing for Migrant Professionals: New Directions for the Occupational English Test. (1986). A report submitted to the Council on Overseas Professional Qualifications by the Testing and Evaluation Consultancy Institute for English Language Education, University of Lancaster, UK. Project members: Alderson, J.C., Candlin, C.N., Clapham, C.M., Martin, D.J., & Weir, C. (Example - David Mendelsohn and Gail Stewart at York University, Ontario were contracted by the Council of the Ontario College of Midwives with provincial funding to develop a language proficiency test - Midwives' English Proficiency Test (MEPT). They have just completed it. David Mendelsohn said that to determine the specifications for the test they consulted widely with people and the document that helped them the most was the above.)

Besides the above references from Canadian University professors, the following were also repeatedly mentioned:

CPA Guidelines... (see record CA02)
APA Guidelines... (see record US01)
ETS Standards... (see record US02) (Example - Stan Jones, Carleton University in Ontario, has done adult literacy surveys for Statistics Canada. He uses ETS Standards and sometimes contracts with ETS for them to do sensitivity or test edit reviews.)

==========

CA01: Canada: "Principles for Fair Student Asmt. Practices..."

Record: CA01
TFTSmem: Turner
Country: Canada
Auth/Pub: Edmonton, Alberta: Joint Advisory Committee
Title(s): Principles for Fair Student Assessment Practices for Education in Canada. (1993)
Language: Available in English and French
Contact: W. Todd Rogers, Chair, Working Group and Joint

Advisory Committee, Centre for Research in Applied Measurement and Evaluation, Faculty of Education, 3-104 Education Centre North, University of Alberta, Edmonton, Alberta, Canada, T6G 2G5. Tel: 403-492-3762 Fax: 403-492-3179

Govt/Priv: priv
Lang/Ed: ed
StanDef: guideline

Comments:

(See also 'Explanatory Notes regarding records from Canada' just before this record.)

Objective/Purpose:
To provide a nation-wide consensus on a set of principles and related guidelines generally accepted by professional organizations as indicative of fair assessment practice within the Canadian educational context. Assessments depend on professional judgment; the objective of this document is to identify the issues to consider in exercising professional judgment and in striving for the fair and equitable assessment of all students.

Target Group/Audience intended:
The information is aimed at all those involved in assessment within the educational context in Canada: both developers and users of assessment. These terms are defined explicitly.

Procedure:
Document developed by a Working Group guided by a Joint Advisory Committee. The latter included two representatives appointed by major Canadian professional organizations: e.g. Canadian School boards Association, Canadian Psychological Association, etc. In addition, the Committee included one representative from each of the Provincial and Territorial Ministries and Departments of Education.

Scope of influence:
The Joint Advisory Committee is seeking to obtain endorsement from all organizations, Ministries, and Departments of Education that are involved in any way in the educational context of Canada. This document is in the process of being disseminated. Forums, etc. are set up for this purpose and suggestions for revision are invited. To date, the majority have provided endorsements. This is an ongoing process at the moment, due to the newness of the document. Not considered exhaustive nor mandatory, but endorsements are looked upon as commitments.

Summary:
This document is not specific to language assessment, but rather encompasses all educational contexts. It is therefore, very applicable. Due to the nature of Canada, that is a bilingual country (French and English) in addition to its diverse cultural

make-up, the document often alludes to the specific language used in assessment and warns that instruments translated into a second language or transferred from another context or location should be accompanied by evidence that inferences based on these instruments are valid for the intended purpose.

It is organized in two parts. Part A is directed at assessments carried out by teachers at the elementary and secondary school levels (also applicable to post-secondary with minor modifications which are specified). Part B is directed at standardized assessments developed external to the classroom by commercial test publishers, provincial and territorial ministries and department of education, and local school jurisdictions. Each Part contains five sections: 1) Developing and choosing methods for assessment, 2) Collecting assessment information, 3) Judging and scoring student performance, 4) Summarizing and interpreting results, and 5) Reporting assessment findings.

Commentary:
This is an exciting document for Canada, because it is the first of its kind, being non-province specific, and being developed by a representative Joint Advisory Committee. Universities are beginning to introduce it in their testing and evaluation courses in teacher training programs. I, myself, have done so in my TESL Testing and Evaluation Course. It motivates much discussion and gives well-founded direction.

==========

CA02: Canada: [CPA] "Guidelines for Ed. and Psych. Testing ..."

Record: CA02
TFTSmem: Turner
Country: Canada
Auth/Pub: Canadian Psychological Association (CPA)
Title(s): Guidelines for Educational and Psychological Testing (1986)
Language: English
Contact: Address: CPA, 151 rue Slater St., Suite 205, Ottawa, Ontario, Canada K1P 5H3. Tel: 613-237-2144 Fax: 613-237-1674
Govt/Priv: priv
Lang/Ed: ed
StanDef: guideline

Comments:

(See also 'Explanatory Notes regarding records from Canada' just before record CA01.)

Objective/Purpose:
To provide criteria for the evaluation of tests, testing practices, and the effects of test use within Canada. These

Guidelines are intended to provide a frame of reference to ensure that relevant issues are addressed by all parties involved in making the professional judgments necessary in testing. They were formulated with the intent of being consistent with the APA Standards (USA), however, due to the differing legal and social contexts within which they operate, CPA preferred to development a statement specifically grounded in the Canadian context. These Guidelines reflect consensus among Canadian professionals and, therefore, supersede the APA Standards in Canada.

Target group/Audience intended:
The information is aimed at all those involved in educational and psychological testing within the context of Canada: at both the individual and institutional levels. Three main participants are identified: test developer, test user, and test taker. A fourth is often involved, that being a test sponsor. These terms are defined explicitly.

Procedure:
The Guidelines were produced by a committee sponsored by CPA. The committee had an advisory group as well as major contributions from professional and academic institutions across Canada. In addition, well-known documents were consulted (e.g. APA Standards, USA).

Scope of influence:
Within its target group, the CPA states, "All professional test developers, sponsors, publishers, and users should make reasonable efforts to observe these Guidelines and to encourage others to do so."(p. 2). Also the CPA recognizes that the use of these Guidelines in litigation is inevitable.

Summary:
As with the APA Standards (see record US01), this document is not specific to language testing, but of course applicable. Due to Canada's bilingual make-up and cultural diversity, however, language considerations are emphasized.

The Guidelines provide a technical guide and basis for evaluating testing practice. A comprehensive glossary is included for term clarification. Guidelines are divided into those of PRIMARY and SECONDARY importance. Since test instrumentation will vary with application, there is a third category designated as CONDITIONAL. The text is divided into three parts. Part I focuses on general guidelines for validity, reliability, test development, comparability, norming, scaling, and publication. Part II presents guidelines for test use. Part III presents administrative procedures. An extensive Bibliography is included.

Commentary:
Several institutions and organizations referred to this document as their only reference. Other than CPA, they use their own expertise and experience.

An example of an influential body using CPA is the Public Service Commission of Canada: Language Service Division. Civil servants are required to be bilingual, and it is this commission that produces the tests for the screening/application process.

==========

CA03: Canada: "Annotated List of French Tests"

Record: CA03
TFTSmem: Turner
Country: Canada
Auth/Pub: Lapkin, S., V. Argue and K.S. Foley
Title(s): "Annotated List of French Tests". Canadian Modern Language Review 48(4), 780-807. (1992)
Language: English
Contact: Sharon Lapkin, Modern Language Centre, Ontario Institute for Studies in Education (OISE), 252 Bloor St. West, Toronto, Ontario, Canada M5S 1V6
Govt/Priv: priv
Lang/Ed: lang
StanDef: test

Comments:

(See also 'Explanatory Notes regarding records from Canada' just before record CA01.)

Objective/Purpose:
To provide educators with a reference guide to French language tests available and in use at the elementary or secondary school level in Canada.

Target group/Audience intended:
All educators and researchers involved in assessing French competency and/or FSL programs and French language programs for francophones (i.e., native speakers).

Procedure:
Authors of the list invited submissions of French tests across Canada. For inclusion in the LIST, three main criteria were used: (1) Are the tests suitable for elementary and/or secondary school students? (2) Do they recur with some regularity in relevant Canadian evaluation literature? and (3) Are they readily available?

Scope of influence:
A reference list referred to and contributed to by a majority of Provincial Ministries of Education (as stated by them in response to our TFTS survey asking for information about standards and guidelines). Also tests are used with some regularity in Canadian educational and research contexts according to the

LIST's inclusion criteria.

Summary:
This article provides an annotated list of French language tests used in education and research in elementary and secondary school contexts with Canada. Twenty-four different tests are described. They are classified according to three types of programs: core French, French immersion, and French mother tongue education. The intended population is specified, but it is mentioned that many tests have been used in more than one type of program. When possible readers are referred to administration guides containing norms. A test index and directory of distributors is provided.

Commentary:
It was interesting to see the amount of provinces/institutions that referred to or contributed to this Annotated List.

==========

CA04: Canada: "ESL Instruction in the Junicr High School ..."

Record: CA04
TFTSmem: Turner
Country: Canada
Auth/Pub: Edmonton, Alberta: Alberta Department of Education, Language Services Branch. 165 pages.
Title(s): ESL Instruction in the Junior High School: Curricular Guidelines and Suggestions. (1988)
Language: English
Contact: Alberta Education, Language Services Branch, Devonian Building, 11160 Jasper Ave, Edmonton, Alberta, Canada TFK 0L2
Govt/Priv: govt
Lang/Ed: lang
StanDef: guideline

Comments:

(See also 'Explanatory Notes regarding records from Canada' just before record CA01.)

Objective/Purpose:
To guide teachers involved in ESL instruction in the junior high schools in order to promote the language development of ESL students in all areas of the provincial curriculum.

Target group/intended:
Educators in the Alberta public school system.

Procedure:
Not specified.

Scope of influence:
Distributed and available to all public junior high schools who have an ESL population.

Summary:
The guide is divided into five sections with the fourth section being pertinent to our TFTS survey: (1) ESL cross-cultural adaption and implications for the classroom teacher, (2) research on the characteristics of language, components of communication, and implications for second language learning, (3) particular needs of adolescent ESL students and how these needs may be met, and (4) suggestions and guidelines for assessment, evaluation, and placement techniques, and examples of reporting procedures, and (5) instructional approaches and classroom techniques.

==========

CA05: Canada: "General Policy for Educational Evaluation..."

Record: CA05
TFTSmem: Turner
Country: Canada
Auth/Pub: Québec, Québec: Direction de l'évaluation pédagogique, Gouvernement du Québec, Ministère de l'Education. (#16-7500)
Title(s): General Policy for Educational Evaluation: for Preschool, Elementary and Secondary Schools. (1981)
Language: Available in French and English
Contact: Gouvernement du Québec, Ministère de l'Éducation, 600, rue Fullum, 9e Étage, Montréal, Québec, Canada H2K 4L1
Govt/Priv: govt
Lang/Ed: ed
StanDef: guideline

Comments:

(See also 'Explanatory Notes regarding records from Canada' just before record CA01.)

Objective/Purpose:
To specify the intentions of the MEQ in the field of educational evaluation: e.g. establish respective roles, assist schools in the development of evaluation policies and practices, promote evaluation that will of greater service to pupils, have evaluation be an integral part of the teaching and learning process, have it be closely related to the curriculum.

Summary:
This document sets out the values, foundation, aims, and basic concepts of educational evaluation within the Québec context. It discusses in detail roles of responsibility and the different

components of a support system in order to pursue a consistent evaluation process: e.g. the place of formative and summative evaluation, the need and means for teacher training in evaluation.

Commentary:
This document continues to be the policy for evaluation practice. Due to increasing decentralization in the province, however, local school boards are presently being asked to develop their own policies specific to their situations, but based on this document.

==========

CA06: Canada: "[Guide to classroom eval.: sec. language...]"

Record: CA06
TFTSmem: Turner
Country: Canada
Auth/Pub: Québec, Québec: Direction de l'évaluation pédagogique, Gouvernement du Québec, Ministère de l'Éducation.
Title(s): Guide d'Évaluation en Classe: Primaire, Langues Seconde, Anglais et Français. (1983) (#16-7220-07) (Title translation: Guide to classroom evaluation: second language at the elementary level, ESL and FSL.)
Language: French, but all ESL examples are in English
Contact: Gouvernement du Québec, Ministère de l'Éducation, 600, rue Fullum, 9e Étage, Montréal, Québec, Canada H2K 4L1
Govt/Priv: govt
Lang/Ed: lang
StanDef: guideline

Comments:

(See also 'Explanatory Notes regarding records from Canada' just before record CA01.)

Objective/Purpose:
To guide all second language elementary teachers (either ESL or FSL) in planning, developing, and implementing formative evaluation procedures for their classrooms that correspond to the provincial curriculum.

Summary:
This document provides a systematic process for classroom formative evaluation. It identifies domains and corresponding program objectives, and then presents specific steps in how to construct appropriate items and tasks. Several prototypes are presented and explained.

==========

CA07: Canada: "[Guide to classroom eval.: formative...]"

Record:	CA07
TFTSmem:	Turner
Country:	Canada
Auth/Pub:	Québec, Québec: Direction de l'évaluation pédagogique, Gouvernement du Québec, Ministère de l'Éducation.
Title(s):	Guide d'Évaluation en Classe: Évaluation Formative, Anglais Language Seconde, Secondaire. (#16-7221-20) (1985). (Title translation: Guide to classroom evaluation: formative evaluation at the secondary level, ESL).
Language:	French, but all items/tasks are in English
Contact:	Gouvernement du Québec, Ministère de l'Éducation, 600, rue Fullum, 9e Étage, Montréal, Québec, Canada H2K 4L1
Govt/Priv:	govt
Lang/Ed:	lang
StanDef:	guideline

Comments:

(See also 'Explanatory Notes regarding records from Canada' just before record CA01.)

Objective/Purpose:
Same as CA06 EXCEPT specific to ESL Secondary Level teachers only.

Summary:
Same as CA06.

==========

CA08: Canada: "[Developing a criterion-referenced...]"

Record:	CA08
TFTSmem:	Turner
Country:	Canada
Auth/Pub:	Québec, Québec: Gouvernement du Québec, Ministère de l'Éducation.
Title(s):	L'Elaboration d'un Instrument de Mesure a Interpretation Criteriee (Perfectionnement en Evaluation Pedagogique). (1986). (Title translation: Developing a criterion-referenced instrument, Professional development in evaluation)
Language:	French
Contact:	Gouvernement du Québec, Ministère de l'Éducation, 600, rue Fullum, 9e Étage, Montréal, Québec, Canada H2K 4L1

Govt/Priv: govt
Lang/Ed: lang
StanDef: guideline

Comments:

(See also 'Explanatory Notes regarding records from Canada' just before record CA01.)

Objective/Purpose:
To provide an approach for developing a criterion-referenced instrument within the public school context and one that corresponds to the provincial curriculum.

Summary:
A step by step process in criterion-referenced instrument development is provided with accompanying activities (this is a general document, not necessarily specific to second language instruction). Integration of the evaluation content into the teaching/learning process is stressed. Table of specification examples are provided focusing on objectives and appropriate item/task types. A procedure for instrument revision and refinement is discussed.

==========

CA09: Canada: "Evaluation, Module 4."

Record: CA09
TFTSmem: Turner
Country: Canada
Auth/Pub: Gascon, L.; Québec: Direction de la formation du personnel scolaire, Gouvernement du Québec, Ministère de l'Éducation.
Title(s): Evaluation, Module 4.
Language: English
Contact: Gouvernement du Québec, Ministère de l'Éducation, 600, rue Fullum, 9e Étage, Montréal, Québec, Canada H2K 4L1
Govt/Priv: govt
Lang/Ed: lang
StanDef: guideline

Comments:

(See also 'Explanatory Notes regarding records from Canada' just before record CA01.)

Objective/Purpose:
To support and guide ESL teachers in their oral production evaluation process within the province's revised curriculum, that is, the communicative approach.

Summary:
Written in a five-day workshop format including Workshop Leader and Participant materials. Covers information such as: the effects of the communicative approach on evaluation/testing procedures, analysis and criteria of items/tasks for both formative and summative evaluation, STANDARDS for instruments, process for development and examples of appropriate tasks.

==========

CA10: Canada: "[Definition of domain for French as a Sec...]"

Record: CA10
TFTSmem: Turner
Country: Canada
Auth/Pub: Direction générale des programmes, Direction de la formation générale des jeunes, Gouvernement du Québec, Ministère de l'Éducation.
Title(s): Definition du Domaine: Français, Langue Seconde. (1988-1992). (Title translation: Definition of domain for French as a Second Language)
Language: French
Contact: Gouvernement du Québec, Ministère de l'Éducation, 600, rue Fullum, 9e Étage, Montréal, Québec, Canada H2K 4L1
Govt/Priv: govt
Lang/Ed: lang
StanDef: performance

Comments:

(See also 'Explanatory Notes regarding records from Canada' just before record CA01.)

Objective/Purpose:
To identify and describe the essential components in the FSL program in preparation for the development of any summative evaluation instrument. To describe the domain so as to clearly establish the relation between teaching the program and testing the program.

Summary:
A series of publications, each one specific to a certain level of FSL instruction in the public school system (e.g. elementary vs. secondary). Program orientation is specified according to sociolinguistic, sociocultural, and pedagogical factors. Corresponding evaluation principles to be followed are provided. These pertain to summative evaluation instruments.

==========

CA11: Canada: "[The effects of the language used in eval. ...]"

Record: CA11
TFTSmem: Turner
Country: Canada
Auth/Pub: Gouvernement du Québec, Ministère de l'Éducation
Title(s): Impact du Choix de la Langue Utilesée dans les Épreuves en Anglais Langue Seconde (1990). [Title translation: The effects of the language used (i.e., English or French) in evaluation instruments for ESL.]
Language: French
Contact: Gouvernement du Québec, Ministère de l'Éducation, 600, rue Fullum, 9e Étage, Montréal, Québec, Canada H2K 4L1
Govt/Priv: govt
Lang/Ed: lang
StanDef: performance

Comments:

(See also 'Explanatory Notes regarding records from Canada' just before record CA01.)

Objectives/Purpose:
To investigate the difference in student performance on ESL provincial exams when instructions and items are presented in either English or French.

Summary:
In Québec, the tradition has been to produce all instructions and items on ESL provincial exams in the native language of the test takers (i.e., French). As demographics change in the province, there has been an interest in changing this policy to the use of only English on these exams to avoid bias for certain linguistic groups. This material is an official document which reports on a Ministry of Education study investigating the use of French and English on provincial ESL exams. No significant differences were found, however, closer examination favored the use of English only. This document recommends that all test instruments for ESL be written and administered in English. It has consequences for future evaluation procedures used for ESL assessment.

==========

CA12: Canada: "[Guide to test construction, ESL, 2nd Cycle...]"

Record: CA12
TFTSmem: Turner
Country: Canada
Auth/Pub: Lord, G. (1990); Québec, Québec: Direction du développement de l'évaluation, Gouvernement du Québec.
Title(s): Guide d'Élaboration d'Instruments de Mesure: Anglais, Langue Seconde, 2e Cycle Secondaire (#16-

7290-02) (Title translation : Guide to test construction, ESL, 2nd Cycle Secondary School).
Language: French, but all item/task examples are in English
Contact: Gouvernement du Québec, Ministère de l'Éducation, 600, rue Fullum, 9e Étage, Montréal, Québec, Canada H2K 4L1
Govt/Priv: govt
Lang/Ed: lang
StanDef: guideline

Comments:

(See also 'Explanatory Notes regarding records from Canada' just before record CA01.)

Objective/Purpose:
To provide guidance to all those involved in constructing, revising, or adapting ESL tests at the final cycle of secondary school in correspondence with the provincial curriculum.

Summary:
Specific procedures are laid out for test construction focusing on the definition of domain, appropriate items/tasks (mainly emphasizing a thematic approach (e.g. storyline), evaluation and revision of the test. Examples are provided.

==========

CA13: Canada: "[Collection of ESL reading test items/tasks...]"

Record: CA13.
TFTSmem: Turner
Country: Canada
Auth/Pub: Lefebvre, D.; Québec, Québec: Gouvernement du Québec.
Title(s): Repertoire des Taches Evaluatives aux Epreuves du MEQ (Juin 1989-1993): Anglais, Langue Seconde, Comprehension d'un Discours Ecrit, 2e Cycle du Secondaire. (1993) (Title translation: Collection of ESL reading test items/tasks from MEQ provincial exams, June 1989-1993)
Language: French, but all items/tasks are in English
Contact: Gouvernement du Québec, Ministère de l'Éducation, 600, rue Fullum, 9e Étage, Montréal, Québec, Canada H2K 4L1
Govt/Priv: govt
Lang/Ed: lang
StanDef: test

Comments:

(See also 'Explanatory Notes regarding records from Canada' just before record CA01.)

Objective/Purpose:
To provide a bank of item/task examples for all those involved in test construction for this final level of ESL in the public school system, and to provide teachers with items/tasks to help students having difficulty

Summary:
Collection of 91 evaluation items/tasks used in provincial exams (secondary leaving exams) from 1989 to 1993 covering all learning objectives of the curriculum.

==========

CA14: Canada: "[Guide to evaluating speaking in class, ESL,...]"

Record: CA14
TFTSmem: Turner
Country: Canada
Auth/Pub: (Government document, but authors Lord, G. & Lord, J.). Québec, Québec: Direction de la formation générale des jeunes.
Title(s): Guide d'Évaluation de la Production d'un Discours Oral en Classe: Anglais, Langue Seconde, 2e Cycle du Secondaire. (1992). (Title translation: Guide to evaluating speaking in class, ESL, 2nd cycle secondary)
Language: French, but all tasks are in English
Contact: Gouvernement du Québec, Ministère de l'Éducation, . 600, rue Fullum, 9e Étage, Montréal, Québec, Canada H2K 4L1
Govt/Priv: govt
Lang/Ed: lang
StanDef: guideline

Comments:

(See also 'Explanatory Notes regarding records from Canada' just before record CA01.)

Objective/Purpose:
To provide guidance to all teachers/officials who administer ESL provincial oral exams at the secondary level. To ease the transition from the traditional individual interview format to a group testing format.

Summary:
In June 1993, the Québec Ministry of Education changed the ESL oral provincial exam format from individual interviews to group discussion. This Guide provides specific information and tools for all of those involved in administering this new type of oral exam. In addition, six activities are provided in order that classroom teachers can engage their students in similar-type

activities during the regular school year.

==========

CA15: Canada: "[Guide to test construction, ESL, 2nd Cycle...]"

Record: CA15
TFTSmem: Turner
Country: Canada
Auth/Pub: Guay, S. & O'Neill (1995), SPEAQ
Title(s): Guide d'Évaluation d'Instrument de Mesure, Anglais, Langue Seconde, 2e Cycle du Primaire (Title translation: Guide to test construction, ESL, 2nd Cycle Primary School).
Language: French, but all item/task examples are in English
Contact: SPEAQ, 7400 boulevard Saint-Laurent, bureau 530, Montréal, Québec, Canada H2R 2Y1.
Govt/Priv: collaboration of govt. and priv.
Lang/Ed: lang
StanDef: guideline

Objective/Purpose:
To provide guidance to all those involved in constructing, revising, or adapting ESL tests at the final level of primary school (i.e., elementary grades 4-6) in correspondence with the provincial curriculum.

Summary:
Specific procedures are laid out for test construction focusing on the definition of domain, appropriate items/tasks (mainly emphasizing a thematic approach, e.g. storyline), evaluation and revision of the test. Numerous examples are provided.

==========

CA16: Canada: "[Guide to test construction, ESL, 1st Cycle...]"

Record: CA16
TFTSmem: Turner
Country: Canada
Auth/Pub: Lord, G. (1995), SPEAQ
Title(s): Guide d'Évaluation d'Instrument de Mesure, Anglais, Langue Seconde, 1er Cycle du Secondaire (Title translation: Guide to test construction, ESL, 1st Cycle Secondary School).
Language: French, but all item/task examples are in English
Contact: SPEAQ, 7400 boulevard Saint-Laurent, bureau 530, Montréal, Québec, Canada H2R 2Y1.
Govt/Priv: collaboration of govt. and priv.
Lang/Ed: lang
StanDef: guideline

Objective/Purpose:

To provide guidance to all those involved in constructing, revising, or adapting ESL tests at the final first level of secondary school in correspondence with the provincial curriculum.

Summary:
Specific procedures are laid out for test construction focusing on the definition of domain, appropriate items/tasks (mainly emphasizing a thematic approach, e.g. storyline), evaluation and revision of the test. Numerous examples are provided.

==========

CA17: Canada: "The Ontario Test of Engl. as a 2nd Lang. ..."

Record: CA17
TFTSmem: Turner
Country: Canada
Auth/Pub:
Title(s): The Ontario Test of English as a Second Language (OTESL): Users' Manual and Final Report (1986).
Language: English
Contact: Marjorie Wesche, Second Language Institute, University of Ottawa, 600 King Edward avenue, Ottawa, Ontario, Canada, K1N 6N5.
Govt/Priv: priv
Lang/Ed: lang
StanDef: test

Comments:

(See also 'Explanatory Notes regarding records from Canada' just before record CA01.)

Objective/Purpose:
To provide specific information for administration, scoring, and interpretation of this performance-based test of English for academic purposes. To provide background information on the test development process.

Target group/Intended audience:
The two documents were designed for all educators/researchers making use of the OTESL battery. The test was originally intended to be used with non-native speakers of English entering Ontario post-secondary programs.

Procedure:
Developed by a team of researchers from Ontario post-secondary institutions under contract to the Ontario Ministry of Colleges and Universities. Project completed in 1989. The test battery was based on a survey of the language needs of ESL students in Ontario post-secondary institutions in combination with

institutional ESL testing needs. Material sources considered and finally used include authentic academic texts and oral discourse. The intent was to develop tasks to measure the real-life language demands faced by ESL students in academic situations.

Scope of influence:
To date the test battery is being used in higher-learning institutions across Canada. Components of it have also been used for research purposes focusing on performance-based testing.

Summary:
(Taken directly from Mari Wesche's communique.)
The OTESL battery consists of three components: (1) Placement Test, 45 minutes, for use in ESL intensive programs with students at lower levels of proficiency, (2) Post Admission Test, 2 1/2 hours, of reading, listening and writing to evaluate the candidates' ability to undertake a full academic program in English, and (3) Oral Interaction Test. All three tests provide diagnostic information and the latter two can be used for purposes of certification. The documents provide the user with much detailed information about procedure and the rationale. It is stated that performance-based testing has special characteristics (e.g. high predictive validity, positive washback effect, high motivational value, and potential for specific, context related diagnostic feedback. It is suggested that these should outweigh the considerations of short-term practicality.

==========

CA18: Canada: "The Canadian Test of Engl... (CanTEST)..."

Record: CA18
TFTSmem: Turner
Country: Canada
Auth/Pub:
Title(s): (1) The Canadian Test of English for Scholars and Trainees (CanTEST): Information Booklet for Test Candidates
(2) A variety of additional Guideline materials: Guidelines for test content for item writers, Guidelines for score users to help in interpreting results, Questionnaires for teachers and examinees to obtain feedback on the test, Description of oral and writing rating scales, A "Technical Report for Etap Version G", Information on equating procedures (Journal article: M. Des Brisay. (1992) Ensuring comparability for CanTEST scores. Monday Morning/Lundi Matin, 5.2, 27-30)
Language: English and French
English version - CanTEST and related documents
French version - TESTCan and related documents
Contact: Margaret Des Brisay, Second Language Institute,

University of Ottawa, 600 King Edward avenue,
Ottawa, Ontario, Canada, K1N 6N5.
Govt/Priv: priv
Lang/Ed: lang
StanDef: test

Comments:

Objective/Purpose:
(1) To provide specific information for test candidates before they come to take the test.
(2) To provide guidelines for administrators, score users, and item writers in each of their respective roles.

Target group/Intended audience:
(1) Non-native English speakers (CanTEST) and non-native French speakers (TESTCan) who want to benefit from either university study or professional exchange in Canada (in English or French respectively).
(2) The various documents were designed for educators/researchers making use of the CanTEST/TESTCan.

Procedure:
The documents above were written and developed at the University of Ottawa. The various versions of the CanTEST and TESTCan (i.e., the topic of the documents above) were developed and validated at the University of Ottawa, under contract with the Canadian/China Language and Cultural Program which is administered by Saint Mary's University in Halifax, Nova Scotia and funded by the Canadian International Development Agency.

Scope of influence:
(Taken directly from a draft article by Margaret Des Brisay entitled "Principles and considerations in the construction of program specific ESL tests: The CanTEST story".)

CanTEST and its French language version, the "Test pour etudiants et stagiaires au Canada" (TESTCan) were originally developed in response to a request from the Canadian International Development Agency for instruments to measure the English or French language skills of candidates from the People's Republic of China who had been selected to come to Canada for either university study or practical attachments. The use of CanTEST later spread to other overseas human resources development projects and their associated language training programs and, for the past several years, CanTEST has been widely used as an in-house ESL admissions test both here at the University of Ottawa and at other Canadian post-secondary institutions.

Summary:
This will summarize the actual CanTEST(TESTCan) which is the topic of all of the documents above.

The CanTEST (TESTCan) has many versions; therefore, it can more

realistically be referred to as a testing system. It is actually a bank of subtests and test procedures from which tests are compiled to meet client needs. The test measures all four skills. An example of one set of versions would be the academic test which takes approximately 3 hours to administer. This includes: a 50-minute listening comprehension test, a 45-minute writing test, a 75-minute reading test, and when specifically requested by the client, a 15-minute oral interview. All materials are taken from authentic documents. Results are reported as band scores (Bands 1-5). Detailed descriptions of the level of performance corresponding to each band are available to guide interpretation of scores. Data collected following operational use in China and Canada were used in standard setting. Practice materials including audio-tapes are available and can be purchased.

==========

CA19: Canada: "English Language Program: Intens. Curr. ..."

Record: CA19
TFTSmem: Turner
Country: Canada
Auth/Pub: Gnida, S., Kirkwold, L., Prior, H., Thomas, W., & Banko, R.
Title(s): English Language Program: Intensive Curriculum. (1992).
Language: English
Contact: Rosalie Banko, English Language Program, Rm. 4-10G University Extension Centre, 8303 - 112 street, 93 University Campus NW, Edmonton, Alberta, Canada T6G 2T4.
Govt/Priv: priv
Lang/Ed: lang
StanDef: guideline

Comments:

(See also 'Explanatory Notes regarding records from Canada' just before record CA01.)

Objective/purpose:
To provide information about the English Language Program concerning program objectives and testing and evaluation procedures.

Target group/Intended audience:
All ESL educators involved in the English Language Program. The students of this program include adults and university-level students.

Procedure:
Not specified.

Scope of influence:
Basically a university-specific document. It is one of the few institutions that has formalized and published such information.

Summary:
This document is specific to ESL instruction. It outlines the background philosophy and objectives of the curriculum. Suggestions are provided for syllabus design. In addition, there is a section on Testing and Evaluation with the following components: (1) A communicative framework, (2) Placement procedures, (3) Formative evaluation, (4) Summative evaluation, (5) Considerations for test design. The importance of valid, reliable, and practical tests is discussed, followed by guidelines to assist educators in test construction.

==========

CA20: Canada: "Carleton Acad. Engl. Lang. Asmt. (CAEL)..."

Record: CA20
TFTSmem: Turner
Country: Canada
Auth/Pub: Fox, J., Carleton University, Ottawa, Ontario
Title(s): Carleton Academic English Language Assessment (CAEL) - A collection of documents and discussion papers dealing with CAEL:
(1) Psychometric Properties of the CAEL Assessment: An Overview of Test Development, Format, and Scoring Procedures
(2) An Examination of Test Methods in the CAEL Assessment
(3) Carleton Academic English Language Assessment: General information for test takers and users.
(4) The Carleton Academic English Language Assessment: Linking Testing and Learning (co-authored: Fox, J. & Pychyl, T.) - Discussion paper and poster presented at LTRC 1991.
Language: English
Contact: Janna Fox, The Centre for Applied Language Studies, 215 Paterson Hall, Carleton University, Ottawa, Ontario, Canada K1S 5B6.
Govt/Priv: priv
Lang/Ed: lang
StanDef: test

Comments:

(See also 'Explanatory Notes regarding records from Canada' just before record CA01.)

Objective/Purpose:

To provide information to the different stakeholders involved in CAEL Assessment concerning the background of test development, the format and scoring procedures, research done on test method effects, and discussion of CAEL use for placement purposes into EAP classes at Carleton University. The actual CAEL Assessment is used as an alternative to the TOEFL or MELAB.

Target group/Intended audience:
(1) The actual CAEL Assessment is targeted for nonnative speakers of English wanting to enter university programs. It tests their English proficiency and establishes their eligibility as well as providing a formula for gradual admission into full-time study.
(2) The collection of documents in this RECORD are designed for test users/educators/researchers who are interested in the CAEL Assessment.

Procedure:
The CAEL Assessment and the other documents above were written and developed at Carleton University. Early in 1988, a group of EAP (English for academic purposes) professors in the Centre for Applied Language Studies (CALS) began to work with professors in the faculties of Science, Engineering, Social Science and Arts to identify actual language performance requirements in introductory, first-year classes. Versions of the CAEL Assessment were piloted with students in several settings. Since 1989, CALS has used the CAEL Assessment and extensive pilot testing has taken place to do item analyses, to calculate raw score conversion factors, and to assess reliability and validity.

Scope of influence:
To date, the CAEL Assessment is used at Carleton University and serves as a testing model for other institutions who wish to link language teaching, testing and learning.

Summary:
This will summarize the actual CAEL Assessment which is the topic of all of the documents above. (Information taken directly from documents.)

The CAEL Assessment provides a mechanism to : (1) identify students who are able to meet the linguistic demands of full-time study, or (2) place students in EAP credit courses while they take one or more courses in their field of study, or (3) identify students who require full-time ESL/EAP in preparation for later studies. The oral assessment is completed prior to the written portion. Oral proficiency is assessed holistically by a trained interviewer. The written portion is an integrated set of language activities. Listening proficiency is assessed in the context of a lecture. Reading proficiency is assessed in the context of academic readings which either introduce or reinforce the topic of the lecture. Writing proficiency is based on an evaluation of the student's response to a short essay question. The results of each of the assessments (i.e., speaking,

listening, reading and writing) is standardized to a band score, which in turn provides a profile of the student's ability to use language for academic purposes.

==========

CH01: China: "A Brief Introduction to the Engl. Lang. Exam..."

Record: CH01
TFTSmem: Alderson
Country: China
Auth/Pub: Liang Yumin
Title(s): A Brief Introduction to the English Language Examinations administered by the National Education Examinations Authority (NEEA)
Language: English
Contact: Liang Yumin, Assistant Researcher, Director of Foreign Language tests Division, National Education Examinations Authority, #30 Yu Quan Road, Beijing, China, 100039
Govt/Priv: govt
Lang/Ed: lang
StanDef: test

Comments:

An overview of the ways in which the NEEA organises and conducts its 11 examinations in English, for roughly 4 million candidates, with details of the examinations. Also sent 10 volumes in Chinese apparently relating to the specifications of the examinations (including word lists), and sample papers.

==========

EU01: Europe/Int'l: "Assoc. Lang. Testers in Europe (ALTE)..."

Record: EU01
TFTSmem: Alderson
Country: Europe / International
Auth/Pub: Association of Language Testers in Europe (ALTE)
Title(s): The ALTE Code of Practice
Language: English
Contact: Dr M Milanovic, Deputy Director, EFL Division, UCLES, 1 Hills Road, Cambridge, CB1 2EU
Govt/Priv: priv
Lang/Ed: lang
StanDef: guideline

Comments:

This document is of considerable significance for ILTA.

Reviewed in Alderson, Clapham and Wall. 1995. Language Test Construction and Evaluation. Cambridge UP.

==========

EU02: Europe/Int'l: "International Baccalaureate Exams..."

Record: EU02
TFTSmem: Huhta
Country: Europe/International
Auth/Pub: International Baccalaureate Organization (IBO)
Title(s): [see 'Comments' below]
Language: English
Contact: International Baccalaureate Examinations Office; Pascal Close; St Mellons; Cardiff; South Glam CF3 0YP; UK
Govt/Priv: priv
Lang/Ed: ed
StanDef: guideline

Comments:

The IB Examinations Office sent copies of certain pages of their procedural manual for administering the tests. They call the manual Vade Mecum, but it appears that parts of the document are not for publication as they contain confidential information. The pages that were sent dealt with instructions to invigilators (arrival time, seat allocation, materials required and allowed, examination stationery required) and regulations about the conduct of the examination (principal's responsibilities, IBO's right to supervise testing centres, malpractice, expulsion from the examination room).

==========

FI01: Finland: "[National certificate: Finnish lang. test...]"

Record: FI01
TFTSmem: Huhta
Country: Finland
Auth/Pub: Language Centre for Finnish Universities, University of Jyväskylä
Title(s): Yleiset kielitutkinnot: suomen kielen testi. Testaajan ohjeet. Perus- ja keskitaso. [National certificate: Finnish language test. Guidelines for the tester. Basic and Intermediate Levels.] (originally 1990, latest version in 1994). About 30 pages + 7 appendices.
Language: Finnish
Contact: Maritta Leinonen, University of Jyväskylä, P.O. Box 35, 40351 Jyväskylä, Finland fax: 358-41-603 521 e-mail: MLEINONE@jyu.fi

Govt/Priv: govt
Lang/Ed: lang
StanDef: guideline

Comments:

Note: Originally, the Finnish language test was a separate test that pre-dated the National certificate system. The test was created in 1990 to test Finnish language (second and foreign) proficiency of foreigners who needed a separate certificate of their skills. The guidelines for testers was created at the same time to help and standardize the marking and assessment, and the writing of certificates. In 1994, the Finnish language test became part of the National certificate system which started that year.

Objective:
Main objectives: To help testers to mark and assess the test-takers products. To standardize marking and assessment. To help them to administer the tests and prepare them for practical things needed in test administration. To help them to write the certificates. Also: To give testers an overall view of the testing process and their role in it. To offer some background information and rationales concerning the tests; to deepen their understanding of the test.

Target group:
The testers who carry out the Finnish test. (They have to supervise the administration of the test, or administer it themselves. They have to mark and assess the products, and to write the certificates.)

Procedure:
The guidelines were produced by the group of teachers and testers who designed the tests (6-8 people). The test designers came originally from the University of Jyväskylä; nowadays the group includes a member from one of the biggest schools teaching Finnish to immigrants and foreigners. The first version was produced in 1990, and it has been updated several times on the basis of experiences of the group and feedback from dozens of testers. Major revisions took place in 1994 when the Finnish test became part of the National certificate system. The guidelines consists of a set of pages not bound together to make updating the various sections easier (i.e. there is no need to update the whole 'book', just the pages where changes have occurred).

Scope of influence:
Before: The Finnish test was an unofficial enterprise that was accepted by a number of schools and institutions (the number increased yearly as the word of its existence spread). It was the first test of Finnish as a Second/Foreign language that was intended for national use.

Now: The National certificate tests are based on a law passed by the Finnish Parliament in 1994. The law states that the Ministry of Education is responsible for the certificate system and sets up a board to supervise the tests. The tests are administered in dozens of schools and other institutions all over the country.

Summary:
Section 1: The National Certificate system is presented and its role in the country's testing and teaching context is outlined. The section also specifies who is eligible to become a tester and the scope of his/her responsibilities.

Section 2: The subtests, test tasks, text types, and the skills tested are briefly introduced.

Section 3: This section deals with the administration of the test. It explains how test-takers enter the test (problems that may arise; what kind of information should be given), and how testers order the tests. Next, the actual administration procedure is presented (e.g. irregular behavior, instructions to test-takers). Finally, there is a checklist on the matters that must be taken care of after the test (e.g. where the materials are to be returned after marking, what to do with extra papers). Also, this part explains the double-marking (20% is double-marked, sometimes more if necessary).

Section 4: This covers the marking and assessment. First, the general principles are presented: e.g. main focus of assessment, criterion-referencing, and conversion tables for the grades. What follows is a rather detailed description of the subtests task by task: at each task, the points that can affect marking or assessment are listed and the number of maximum points is given, as well as how much should be deducted for various types of errors. These include the rating criteria for writing and speaking, and the weights assigned to certain criteria.

Section 5: This deals with writing certificates (the tester writes them him/herself after allowing time for possible double-marking). (Note: The certificate reports the overall grade, plus separate grades for reading, speaking, listening, writing, and grammar & vocabulary. Short descriptions of what the numbers mean are also provided, as well as the overall proficiency scale of the National certificate system. In the Finnish test the testers are free to write the descriptions; in the other tests the descriptions usually follow more or less automatically from the number-grades.) The guidelines list some points that may cause the tester to modify the standard descriptions that appear in the sample certificates and in the descriptions found in the proficiency scales.

Section 6: The names, addresses, telephone numbers etc. of the test design group are given in case the tester has problems he or she wants to discuss.

Appendices include the proficiency scales used in the system, lists of topics, functions, and structures that the test tasks are based on, various forms, and information to be passed on to the test takers to help them to prepare for the test. In 1995, a video cassette presenting benchmark examples of the oral proficiency levels covered by the basic and intermediate test was added to the 'Guidelines'.

Commentary:
The testers also participate in training: the 'Guidelines for testers' is sent to them in advance, and is discussed during the training.

==========

FI02: Finland: "[National certificate: test specifications...]"

Record: FI02
TFTSmem: Huhta
Country: Finland
Auth/Pub: Language Centre for Finnish Universities, Univ. of Jyväskylä
Title(s): a) Yleiset kielitutkinnot: Ohjeet tehtavien laatimiseksi. Perus- ja keskitaso [National certificate: test specifications for the basic and intermediate level tests] (originally 1993, latest version No.9, May 31, 1995). 47 pp.
b) Yleiset kielitutkinnot: Ohjeet tehtavien laatimiseksi. Ylin taso [National certificate: test specifications for the advanced level tests] (originally 1993, latest version No.9, June 6, 1995). 45 pp.
Language: Finnish
Contact: Maritta Leinonen, University of Jyväskylä, P.O. Box 35, 40351 Jyväskylä, Finland fax: 358-41-603 521 e mail: MLEINONE@jyu.fi
Govt/Priv: govt
Lang/Ed: lang
StanDef: test

Comments:

Note: National certificate = The Finnish National Certificate of Language Proficiency. This is a testing and certification system for mainly adult learners of foreign languages who need a certificate for study or work, or who are interested in their own progress. There are three levels: Basic, Intermediate, and Advanced. The languages in the National certificate system are English, Swedish, German, French, Russian, Spanish, and Finnish (as a second and foreign language).

Objective:

To help test designers produce comparable tests in different languages, and different versions in the same language; to standardize test design (text selection, task types, questions formats). In other words, to help them translate the proficiency scales and lists of topics, subjects and functions into test tasks (these scales and lists are presented in another publication which is also available to the public, e.g. test-takers).

Target group:
Test designers and item writers in the languages included in the National certificate system

Procedure:
One or two members of the test design group produced draft specifications which were then revised on the basis of discussion in the group and feedback from a supervisory group. The main test design group consisted of research and teaching staff at the University of Jyväskylä (8-10 persons). The supervisory group consisted of two test designers from the university and a number of teachers and administrators representing the National Board of Education and various schools and institutions of adult education. The specifications follow, with some modifications, the format presented by Davidson, F. and Lynch, B. (1993) in their article Criterion-Referenced Language Test Development: A Prolegomenon (In A. Huhta, K. Sajavaara, and S. Takala (Eds.) Language Testing: New Openings (pp. 73-89). University of Jyväskylä, Institute for Educational Research.)

Scope of influence:
The National certificate tests are based on a law passed by the Finnish Parliament in 1994. The law states that the Ministry of Education is responsible for the certificate system and sets up a board to supervise the tests. The tests are administered in dozens of schools and other institutions all over the country.

Summary:
The specifications present the subtests one by one and specify for each the following points:

a) Explanation of the specific terminology used
b) General and specific description of the skills tested (short description of the behaviour to be tested and the purpose of the test)
c) Description of and the requirements for the test tasks
   - what the test-taker will encounter
   - how the rubric and instruction should be presented
   - content, topic
   - difficulty
   - format
   - type of text or discourse, functions
d) What the test-taker is expected to perform
e) Marking or assessing the responses
f) Sample tasks

g) Requirements for the texts
- content, topic
- difficulty
- format
- type of text or discourse, functions

h) Additional information
The specifications also include information as to where and how the tests for different languages are assumed or allowed to differ from each other.

==========

FI03: Finland: "The Finnish Matriculation Examination"

Record: FI03
TFTSmem: Huhta
Country: Finland
Auth/Pub: The Matriculation examination board
Title(s): The Finnish Matriculation Examination
Language: Finnish
Contact: The Matriculation examination board; Helsinginkatu 34 C; 00530 Helsinki; Finland
Govt/Priv: govt
Lang/Ed: ed
StanDef: test

Comments:

Finland has a school-leaving examination that is comparable to the British A-levels or the French baccalaureate, called the Matriculation examination. The exam covers various domains of knowledge, including foreign languages. Only the students who are in the academically oriented senior secondary school take this examination. The Matriculation examination is designed by a special Board set up for the purpose. As in the Swedish 'centrala prov', there are no detailed, written guidelines available for test construction, although general guidelines as to e.g. what test types are to be used exist, and are available to teachers at schools. Since the tests are first marked by the school teachers, there are some documents that guide the marking, and the administration of the tests at schools. The second marking / rating is done centrally by the members and external markers contracted for the purpose by the Board. The following information about the foreign language tests is given to the teachers in various brief documents for each examination (twice a year):

a) The right answers for the multiple-choice reading and listening comprehension and vocabulary and grammar tests. The correct and partially correct answers for the short answer questions found in some sections of the comprehension tests.

b) The assessment criteria and a collection of benchmark compositions for rating the essay writing tasks.

c) The schools also receive a general guideline specifying the administration of the examination, as well as general descriptions of, e.g. the test types used.

d) The external assessors receive more detailed assessment guides which in addition to document b) also give advice on various practical problems and questions that an assessor might encounter.

The Matriculation examination is adopting new more open-ended test formats, which means that additional guidelines will be needed to standardise teachers' marking and rating at schools in the future.

==========

FR01: France: "DELF (Diplôme d'Etudes en Langue Française)..."

Record: FR01
TFTSmem: Huhta
Country: France
Auth/Pub: Commission Nationale du DELF et du DALF
Title(s): DELF (Diplôme d'Études en Langue Française) Guide de l'examinateur. Didier / Hatier 1993. (32 pages)
Language: French
Contact: The document can be ordered from:

Les Éditions Hatier
8, rue d'Assas
75278 Paris cedex 06
FRANCE

Information about the DELF and DALF examinations can be obtained from:

Commission Nationale du DELF et du DALF
Centre International d'Études Pédagogiques
Service des Certifications
en Français Langue Étrangère
1, avenue Leon Journault
92311 SEVRES CEDEX
FRANCE

Govt/Priv: priv
Lang/Ed: lang
StanDef: test

Comments:

The DELF examination is a test of proficiency in French for speakers whose mother tongue is not French. It is part of a

testing system created and supervised by the Ministry of Education of France. The examinations are designed, administered and marked in the country organising the examination. The 'Commission Nationale' in France, which is the body actually responsible for the exams, supervises the testing centres and their tests all over the world to ensure the comparability of the examinations in different countries. There is also a higher level examination in this system (DALF, Diplôme Approfondi de Language Française).

Objectives and target groups:
For (local) assessors to standardise the rating of the candidates in different testing centres and in different countries.

Procedure:
Produced by the Commission Nationale; based on the study of rated candidate papers and performances from the local testing centres.

Scope of influence:
Concerns the assessors at the testing centres for DELF in about 80 countries all over the world. Probably has some influence on the assessment procedures as far as the teaching of French in those countries is concerned.

Summary:
The document first describes, and makes a distinction between, communicative and linguistic competencies, then explains the notion of progression or levels of proficiency underlying the testing system and the assessments. Then, general suggestions are given as to the successful evaluation (e.g. that double marking should be used in productive tests), and certain factors affecting assessment are briefly presented (e.g. halo effect, excessive severity & leniency, letting performance in e.g. classroom to affect exam grades).

Most of the document is devoted to giving guidance on the assessment of the five parts of the examination. For each part, the document specifies the assessment criteria in terms of what the candidate is expected to do with the language (savoir-faire, communicative competence) and in terms of linguistic competence (e.g. which grammatical forms or what kind of lexis should be mastered). This section also includes numerous examples of forms or grids that the assessors should use when rating the candidates. These grids list the criteria to be assessed and the weights (maximum points) for each criterion.

==========

FR02: France: "1994. Guide du concepteur de sujets. DELF-DALF"

Record: FR02
TFTSmem: Huhta
Country: France

Auth/Pub: Dayez, Y.
Title(s): 1994. Guide du concepteur de sujets. DELF - DALF. Paris: Hatier /Didier. (160 pp.)
Language: French
Contact: Les Éditions Hatier; 8, rue d'Assas; 75278 Paris cedex 06; FRANCE. Information about the DEFL and DALF examinations can be obtained from: Commission Nationale du DELF et du DALF; Centre International d'Études Pédagogiques; Service des Certifications en Français Langue Étrangère; 1, avenue Leon Journault; 92311 SEVRES CEDEX; FRANCE
Govt/Priv: priv
Lang/Ed: lang
StanDef: test

Comments:

See also the comments in the summary of 'DELF Guide de l'Examinateur' (record FR01).

This book is a very systematic and thorough guide on designing tests for this testing system, and it combines general specifications with concrete examples of test tasks produced on the basis of these specifications.

Objectives and target group:
The book is intended for the test designers and item writers who produce the DELF and DALF tests at the local testing centres in various countries. The aim is to ensure that the tests designed in different places are comparable enough and follow the guidelines set for the examination system, while allowing certain amount of freedom and creativity in test design. The book should be used together with the 'DELF Guide de l'Examinateur'.

Process:
The book was written by Y. Dayez, apparently representing the Commission Nationale du DELF et du DALF, with the assistance of local testing centres.

Scope of influence:
The test designers and item writers for the two tests in different countries; possibly teachers and teaching of French in the countries where local testing centres are situated.

Summary:
The book begins with a few pages of general remarks concerning the DELF and DALF tests (parts, objectives, levels, testing times, nature of topics and texts, test instructions, and assessment). This includes a checklist of points that can be used to verify the adequacy of a prospective topic or text for the tests. The rest of the book contains more detailed specifications for the two tests, especially for the DELF. The specifications are presented in a clear, structured fashion: for each part (unit) of the test, and each subtest within a unit, certain matters are

specified. These include the general objective of the unit or the subtest, skills tested, choice of topics or texts (e.g. length, difficulty, sources), questions (e.g. content, form, number), instructions to the testee, and marking or assessment. These specifications are accompanied with a varying number of examples - often several - of actual test tasks appropriate for the tests, i.e., texts (written and oral) or topics, and questions and correct answers with marking instructions.

The book also contains brief specifications for creating a test for checking whether a candidate who has not take the DELF has a sufficient command to enter the DALF test.

==========

FR03: France: "Monitoring Education-For-All Goals..."

Record: FR03
TFTSmem: Alderson
Country: France
Auth/Pub: UNESCO
Title(s): Monitoring Education-For-All Goals: A Joint UNESCO-UNICEF Project. March 1994
Language: English
Contact: Prof V Chinapah, Unit for InterAgency Cooperation in Basic Education, UNESCO, 7 Place de Fonteroy, 75780 Paris, France
Govt/Priv: priv
Lang/Ed: ed
StanDef: performance

Comments:

Reports on a Monitoring Project focusing on learning achievement which has been implemented in China, Jordan, Mali, Mauritius and Morocco and which is in the course of implementation in Oman, Sri Lanka, India, Brazil, Ecuador, Nigeria, Tanzania, Lebanon, Mozambique and Slovakia. One of the objectives of the Project is to develop a set of measurable indicators geared to the principle of education-for-all goals (relating to achievement in literacy, numeracy, life skills) and factors affecting achievement (student, home and school characteristics). The participating countries are asked to prepare their own instruments for evaluating learning achievement at basic education level, and this publication reports progress.

The Progress Report contains a description of the major outcomes to date, and selected findings, and appends a brief content analysis of the various tests and questionnaires developed in the participating countries. Guidance was given on report writing schedules and formats, and training of national teams was undertaken by the Project (eg an

International Workshop on Survey Methodologies, a workshop in using SPSS to clean and analyse data). UNESCO supplied prototype questions, but countries were free to deviate from these. The report contains no details of the standards of test construction that might have had to be met and it is not clear whether such standards were in fact produced. The report does give some information about the common core of basic learning competencies, the use of questionnaires, the application of the sampling frame and coverage and the use of an analytical framework for data analysis and report writing. Instruments are available in original and translated versions (some in English, others in French). A future International Progress Report (not received) "will eventually cover the technicality and measurement properties of this project." Preliminary results only are presented in this Report.

Of interest to TFTS is the following: "The application of a common framework agreed on by all participating countries and adjusted to the national contexts helped to ensure validity and reliability in this international project. 'Educational standards' must be regarded as fundamentally 'relative' (Beeby, 1969) A proposed international study should display sensitivity to the cultural contexts, ie language spoken, taught and examined, religion, laws, implements used, values for the education dimensions to be assessed (Bradburn/Gilford, 1992). Measurement features lending to cultural bias should be avoided.....The Monitoring Project sought to avoid common problems of dependency and encourage a participatory approach by:

1) ensuring that all issues relating to the overall project design (i.e. target groups, instrument construction, sampling procedures, data collection, analysis and reporting) were initiated, discussed, pre-tested and fine-tuned by a core group of national experts, under the guidance of the national task force in each country;
2) recognizing the uniqueness of each country's sociocultural, linguistic, developmental and educational characteristics, thus facilitating the analysis of country-based data; and
3) ensuring that the measurement indicators were set, defined and reported by the countries themselves and that basic learning competencies were defined and standards for literacy, numeracy and life-skills set by the countries themselves."

It will be valuable for ILTA to obtain further Reports as they are produced. This Project is probably as significant as the IAEA Language Education Survey (see documents NE02, NE03) for the mission of the TFTS.

==========

GE01: Germany: "[Administering the Goethe-Institut tests...]"

Record: GE01
TFTSmem: Huhta
Country: Germany
Auth/Pub: Goethe-Institut
Title(s): Hinweise zur Durchfürung der Prüfungen des Goethe-Instituts. Checkliste für Prüfer. [Administering the Goethe-Institut tests. Checklists for the tester]. May 1993. Goethe-Institut München. (129 pages)
Language: German
Contact: Contact person: Dr. Jutta Weisz, tel. 089-15921-371 Goethe-Institut, Referat 33 und 43, Helene-Weber-Allee 1, D-806037 Munich, Germany
Govt/Priv: priv
Lang/Ed: lang
StanDef: test

Comments:

The Goethe-Institut's tests are probably the best known tests of German that are produced in Germany.

Objective:
To give information about all aspects of organizing, and administering the Goethe-Institut's tests.

Target group:
All those involved in administering any of the Goethe-Institut's tests (as stated in the document). (On the basis of the content, I would say that it is mainly for those responsible for organizing the tests at the local centre, who are responsible for all the practical details: ordering the material, informing the candidates and testers. Also testers probably find this useful but at least for marking / rating the KDS and GDS tests there is another document that is more informative in that respect.)

Procedure:
The guidebook was developed by the Goethe-Institut on the basis of earlier guidebooks and experience and feedback from tester training seminars.

Scope of influence:
The tests: the tests are internationally known and administered by the Goethe-Instituts in dozens of countries. They apparently are officially recognized at least in Germany.

This document: used at the Goethe-Instituts worldwide.

Summary:
The guidebook contains general information about administering

the Goethe-Institut's tests as well as more specific information about the content and the administration of the tests:

Zertifikat Deutsch als Fremdsprache, ZDaF ("Certificate of German as a Foreign Language" (these translations are entirely unoffical))
- tests a basic-level proficiency (requires about 400-600 hours of intensive learning of German, according to the guidebooks)

Zentralen Mittelstufenpruefung, ZMP ("Central test at the level of secondary education")
- requires a good command of general German (800-1000 hours intensive training); some universities accept this as a proof of German proficiency, if the pass grade is good enough

Pruefung Wirtschaftsdeutsch international, PWD ("International test of German for economic purposes")
- LSP-test of the language needed to do business in German; otherwise the level of difficulty is similar to the ZMP

Zentralen Oberstufenprüfung, ZOP ("Central test for upper secondary education")
- requires good or very good command of German (about 1200 hours)

Kleines Deutsches Sprachdiplom, KDS ("Lesser German language diploma")
- about the same level as ZOP; accepted by the universities as foreigners' certificate of proficiency in German

Grosses Deutsches Sprachdiplom, GDS ("Greater German language diploma")
- requires a near-native proficiency in German

First, some general points concerning each test are listed: contact person and address, availability of practice material, list of publications (guidelines, sets of rules etc) about the test, material for tester training, and some additional points about what kind of information the testing center should give to the candidates, how to organize the test as well as some advice on tester behaviour on specific parts of the test. (about 15 pages)

Then follow several order forms for the tests and various kinds of practice material for them, as well as forms dealing with exceptions in the test fees. (about 25 pages)

The rest of the guidebook contains general information about the above mentioned tests and their administration, presented in a systematic fashion. The matters covered for each test include the following (there is some variation in coverage between the tests):
- goal of the test
- who can enter the test

- where the test is administered and when
- who constitute the testing group
- fees
- how the test is advertised
- exclusion from the test (e.g. due to cheating)
- what constitutes a 'pass' in the test
- retaking the test
- the certificate
- how to complain about the test result

These points are followed (for each test) by a description of the parts of the test (including the times allotted).
This is followed by general guidelines on marking and rating: tables for converting points to grades, listing of maximum points for each task, describing what the tester should do at each point of the test, and some sample letters with original assessor marks, comments and grades. For some tests, rating scales (with verbal descriptions) for the writing and/or speaking tests are also included.

(All in all, this document is more general and contains far fewer examples that the other Goethe-Institut document referred to in record GE02.)

==========

GE02: Germany: "[Lesser German Language Diploma. Greater German...]"

Record: GE02
TFTSmem: Huhta
Country: Germany
Auth/Pub: Goethe-Institut
Title(s): Kleines Deutsches Sprachdiplom. Grosses Deutsches Sprachdiplom. Informationen für Lehrer und Prüfer [Lesser German Language Diploma. Greater German Language Diploma. Information for the teacher and the tester]. (1993) and Hinweise zur Durchfürung der Prüfungen des Goethe-Instituts. Checkliste fuer Prüfer [Administering the Goethe-Institut tests. Checklists for the tester]. (May 1993) (128 pages)
Language: German
Contact: Dr. Sibylle Bolton, tel. 089-15921-382 and Mrs. Jutta Steiff, tel. 089-15921-503 Goethe-Institut, Referat 33 und 43, Helene-Weber-Allee 1, D-806037 Munich, Germany
Govt/Priv: priv
Lang/Ed: lang
StanDef: test

Comments:

The document is a thorough and systematic guide for markers, with extensive samples of test takers' performance.

Objective:
To help testers to mark and assess the test-takers products.

Target group:
Testers (according to the title also teachers who intend to give preparatory courses for the exam.)

Procedure:
These guidelines are produced by the Goethe-Institut, apparently by their test designers and administration officers.

Scope of influence:
The tests: A government-level organ (Kultusministerkonferenz) has recognized that the KDS test is an acceptable proof of a foreigner's German language skill for university studies in Germany (a foreigner with the KDS certificate need not take the otherwise compulsory language test for prospective university students). The GDS test serves the same purpose although it greatly exceeds the language requirements for university study. The GDS certificate indicates a near-native proficiency and is thus widely recognized by private and public organizations all over the world. In some countries, the GDS is recognized as a sufficient proof of the language proficiency of graduating language teachers.

The guidelines: the guidelines probably affect most the teachers who in different countries teach for the tests, since it provides extensive information about the tests and their marking. Probably the guidelines are also used in tester training and for maintaining quality of marking (marking is centralized and takes place in Munich at the local Goethe-Institut and the U. of Munich). Since the speaking test is face-to-face, the assessment apparently takes place on the spot, and thus the interlocutor-testers probably use the guidelines as well.

Summary:
The guidelines consist of a brief general introduction and detailed descriptions of the two tests with extensive examples and advice on how to mark and rate the subtests in general and the samples in particular. Appended are a) some rules governing the entrance to and administration of the tests, and b) list of the international test centers for the two tests.
The emphasis of the document is on the detailed presentation of typical test tasks and guidelines for marking and assessing them.

The following is a translation of the table of contents

1. General information
1.1. Comparison to the other German tests

2. Kleines Deutsches Sprachdiplom (starting at page 4)
2.1. Description of the aims of the test
2.2. Endorsement of the test

2.3. Sample test batteries 1 and 2 (page 9)
2.3.1. Parts of the test and the book list
2.3.2. Sample battery 1
2.3.3. Sample battery 2

2.4. Assessment (page 30)
2.4.1. Giving points and calculating the grades
2.4.2. Assessment criteria for the oral test
2.4.3. Marking the "text with questions about the content and vocabulary" and "tasks on expression ability"/ Sample battery 1
2.4.4. Marking the "text with questions about the content and vocabulary" and "tasks on expression ability"/ Sample battery 2
2.4.5. Marking the "lecture" and the essay

3. Grosses Deutsches Sprachdiplom (page 58)
3.1. Description of the aims of the test
3.2. Endorsement of the test

3.3. Sample test batteries 1 and 2 (page 67)
3.3.1. Parts of the test and the book list
3.3.2. Sample battery 1
3.3.3. Sample battery 2

3.4. Assessment (page 85)
3.4.1. Giving points and calculating the grades
3.4.2. Assessment criteria for the oral test
3.4.3. Marking the "text with questions about the content, vocabulary, and style" and "tasks on expression ability" / Sample battery 1
3.4.4. Marking the "text with questions about the content, vocabulary, and style" and "tasks on expression ability" / Sample battery 2
3.4.5. Marking the "questions on subject matter / professional field", "question on Landeskunde", and the essay.

Appendices (page 114)
A. Administration guidelines for the Kleines Deutsches Sprachdiplom and the Grosses Deutsches Sprachdiplom
B. List of testing centers

==========

HK01: Hong Kong: "The Work of the H.K. Exam. Authority..."

Record: HK01
TFTSmem: Alderson

Country: Hong Kong
Auth/Pub: Hong Kong Examinations Authority
Title(s): The Work of the Hong Kong Examinations Authority, 1977-93
Language: English
Contact: Contact: Rex King, Deputy Secretary, Hong Kong Examinations Authority, Southorn Centre, 13th Floor, 130 Hennessy Road, Wanchai, Hong Kong Fax (852) 2572-9167
Govt/Priv: other (The HKEA is neither government nor private. It is public but receives no money from Government for recurring expenditures.)
Lang/Ed: ed
StanDef: guideline

Comments:

Describes the work of HKEA, with sections on the structure of HKEA and its Committees, how examinations are constructed, markers recruited, feedback given to schools, how examinations are administered and results issued, including those external proficiency tests that HKEA administers on an agency basis. It describes the Systems and Statistics Section of HKEA and the work they do, including examination processing and analysis, information processing and the compilation of statistics, marking and grading procedures. Sample item analyses and correlation matrices are reported. There is a section on The Backwash Effect of Examinations on Teaching, and an appendix dealing with Common Misconceptions about Public Examinations in Hong Kong.

This is an excellent document that would repay detailed study and analysis.

==========

HK02: Hong Kong: "An Introduction to Educational Assessment..."

Record: HK02
TFTSmem: Alderson
Country: Hong Kong
Auth/Pub: Paul Lee and Charles Law, Education Department, Hong Kong
Title(s): An Introduction to Educational Assessment for Teachers in Hong Kong
Language: English
Contact: Contact: HHA Poon, Educational Research Section, Education Dept, 11/F Wu Chung House, 197-221 Queen's Road East, Wanchai, Hong Kong
Govt/Priv: govt
Lang/Ed: ed
StanDef: guideline

Comments:

Responded that they "have no specific publications on "assessment standards" but do provide specific testing instructions and reference materials including assessment exemplars to the test setters/item writers". Enclosed the document above:

Intended for use in initial or in-service teacher training, the book gives guidance on test construction and the uses of assessment, and is divided into chapters covering topics like: Assessment Plans, Assessment Methods, Reporting measurement Results, The Criteria for Judging a Test- Reliability and Validity, Course/Teacher Evaluation Questionnaire, Assessment and Allocation Systems in Hong Kong and Public Examinations in Hong Kong. It contains an initial test in testing concepts and numerous exercises throughout the text. It refers to a publication by the IAEA, entitled 'A Teacher's Guide to Assessment'.

==========

HK03: Hong Korg: "General Introduction to Targets..."

Record:	HK03
TFTSmem:	Alderson
Country:	Hong Kong
Auth/Pub:	Institute of Language in Education (Director: John Clark)
Title(s):	General Introduction to Targets and Target-Related Assessment. First Draft Version Draft TTRA Assessment Guidelines for Subject-Specific Development Groups (1993) Revised Learning Targets for English (1993) TOC (ex-TTRA) Cross-Curricular Framework of Commo... Concepts Draft Bands of Performance for English, 1994
Language:	English
Contact:	Contact: Dr J J Clark, Institute of Language in Education, 2 Hospital Road, Hong Kong Fax (852) 2559 5303
Govt/Priv:	priv
Lang/Ed:	lang
StanDef:	performance

Comments:

In addition, Dr Clark sent three of his own papers: Reflections on Grade-related Criteria, Modern Languages in Scotland; Targets and Target-Related Assessment: Hong Kong's project in educational standards-setting for the improvement of student learning (published in Education Standards for the 21st Century: Papers presented at the Asian Pacific Economic Cooperation Education Symposium, in Washington DC, August 1992: US Department of Education; The Challenge of Standard-Setting: Hong Kong's Target

Oriented Curriculum Scheme. These relate to projects in which he was involved in Graded Levels of Achievement in Foreign Language Learning (GLAFLL), in Scotland; The Australian Language Levels Project (ALL); and the Target Oriented Curriculum Initiative (TOC - formerly TTRA) in Hong Kong. He provides further contacts and addresses for these projects, but they have not been followed up, as they seem tangential to ILTA interests.

These appear to relate to levels of performance, not criteria for evaluating tests.

==========

HK04: Hong Kong: "Public Examinations in H.K., 1993"

Record:	HK04
TFTSmem:	Alderson
Country:	Hong Kong
Auth/Pub:	Hong Kong Examinations Authority
Title(s):	Public Examinations in Hong Kong, 1993 (8 pages)
Language:	English
Contact:	Contact: Rex King, Deputy Secretary, Hong Kong Examinations Authority, Southorn Centre, 13th Floor, 130 Hennessy Road, Wanchai, Hong Kong Fax (852) 2572-9167
Govt/Priv:	other (see HK01 for details)
Lang/Ed:	ed
StanDef:	other

Comments:

Describes the examinations available in Hong Kong, and gives brief details of results and comparabilities, sometimes with TOEFL scores.

Statistical appendices include details of grades awarded, and analysis of age distribution of candidates by sex.

A detailed letter from the Deputy Secretary, in which he reinterprets "your request in more general terms as a desire to know how we ensure the reliability and validity of our system. Here I am using these terms in the broadest sense where the concern is to produce subject grades that make sense across subjects and between years, and where the grades we award are a reliable measure of the things we claim we are measuring". The letter goes on to give a general overview of how reliability and validity are ensured, and offers further information and publications, yet to be requested. It is a very useful contribution.

==========

HK05: Hong Kong: "Statistics Used in Public Examinations..."

Record: HK05
TFTSmem: Alderson
Country: Hong Kong
Auth/Pub: Hong Kong Examinations Authority
Title(s): Statistics Used in Public Examinations in Hong Kong. May 1988
Language: English
Contact: Rex King, Deputy Secretary, Hong Kong Examinations Authority, Southorn Centre, 13th Floor, 130 Hennessy Road, Wan Chai, Hong Kong Fax (852) 2572-9167
Govt/Priv: other (see HK01 for details)
Lang/Ed: ed
StanDef: guideline

Comments:

A very useful 64-page document describing the main statistics used for analysing public tests in Hong Kong. It was written in the hope that "users of the reports from the computer systems will find it useful."

Statistics described, explained, and whose use is exemplified, include:
1. Normal distribution, standard deviation and standard error
2. Question analysis
3. Correlation between m.c. and conventional papers
4. Multiple-choice item analysis
5. Markers' percentile graphs
6. Adjustment of raw marks
7. Combination of paper marks, planned and effective weights
8. Grading system
9. Some commonly-used formulae
10. Chinese translation of some commonly-used terms.

The intended readership for this document are the HKEA professional staff, in particular subject officers; i.e., this is an internal document.

This document is of considerable interest for TFTS and ILTA.

==========

IN01: India: "Handbook of Evaluation in English..."

Record: IN01
TFTSmem: Alderson
Country: India
Auth/Pub: M Agrawal
Title(s): Handbook of Evaluation in English. National Council of Educational research and Training, 1987
Language: English
Contact: Dr M Agrawal, National Council for Educational Research and Training, Sri Aurobindo Marg, New

Delhi 110016, India
Govt/Priv: govt
Lang/Ed: lang
StanDef: guideline

Comments:

A textbook (108 pages) for teachers on English language testing, supplied to TFTS to give "an idea how the paper setters are trained in item writing and paper setting." Contains sections on: Evaluation in English; Instructional Objectives of English; Writing Different Forms of Questions; Testing Elements of Language; Testing Comprehension; Testing Expression; Preparing a Balanced Question Paper and Unit Tests; Appraising the Quality of a Test; Analysing, Interpreting and Using the Test Results; and a sample question paper and bibliography.

In an accompanying letter, Dr Agrawal describes the testing situation in India, and argues for a more communicative approach to testing in India. She says that "The School Boards of Education provide the paper setter with a general design of the paper which indicates the content to be tested and the types of questions to be used. Efforts are made to get the paper setters trained in the technique of paper setting so as to improve the quality of questions. Some Boards even do a post-administration qualitative analysis of the question papers in order to remove the defects, if any, in the future examinations. But still a lot remains to be done specially with regard to the quality of questions."

==========

IR01: Ireland: [NCCA] "Junior Certificate..., Leaflet..."

Record: IR01
TFTSmem: Alderson
Country: Ireland
Auth/Pub: National Council for Curriculum and Assessment
Title(s): i) The junior Certificate 1992: A guide for parents
ii) Leaflet on the NCCA
Language: English
Contact: National Council for Curriculum and Assessment, Dublin Castle, Dublin 2, Ireland. Fax 01 679 8360
Govt/Priv: govt
Lang/Ed: ed
StanDef: test

Comments:

Document i) is a brief description of a major new test for pupils aged about 15 in a range of subjects. No information is available about its syllabus, construction or standards, but

other documents may describe these.

Document ii) briefly describes the work of the NCCA, whose function is to advise the Minister of Education on curriculum, the assessment of pupil progress, on in-service training for teachers, and on "the standards reached by pupils in the public examinations." Current main concerns appear to be syllabus revision and a review of Leaving Certificates, presumably in terms of suitability of content. Further information may be available on request.

==========

MA01: Mauritius: "The Certificate of Primary Education..."

Record: MA01
TFTSmem: Alderson
Country: Mauritius
Auth/Pub: Mauritius Examinations Syndicate
Title(s): The Certificate of Primary Education Examination: Regulations and Syllabuses, 1980, 1987, 1991-92, 1994
Syllabus of English Language, 1996
Syllabuses for Primary Schools Standards I - VI 1985
Learning Competencies for All
PSLC Syllabus 1978
New French Syllabuses at School Certificate and Higher School Certificate Level (both Cambridge examinations)
Language: English
Contact: Contact: R Manrakhan, Principal Research and Development Officer, Mauritius Examinations Syndicate, Reduit, Mauritius. Fax: (230) 4547675
Govt/Priv: govt
Lang/Ed: lang
StanDef: test

Comments:

A number of documents were sent, with a brief overview of the history of examinations in Mauritius, and the suggestion that ILTA contact Moray House, the London Board, and UCLES for further information.

Most of these documents cover the objectives, content and method of the syllabus, include reference to timetables, but also constitute test specifications. In addition, examination regulations are included, covering topics like Entry Requirements, Conditions for the Award of a Certificate, Grading, Ranking (which seems to relate to weighting of papers), exam dates, procedures for replacing lost certificates, disqualifications of candidates, appeal procedures, and other administrative matters.

Learning Competencies for All is a lengthy document which gives considerable detail of the appropriate levels for learning of all school subjects which "not only provides the necessary direction for the improvement of standards of performance, but also more concrete criteria for the certification of achievement". The intention is to encourage competency-based testing, and a "greater element of accountability in the system". The competencies are to be used for revision of examinations, and specimen exemplars of items and tests will be developed. "A criterion-referenced approach to testing is to be adopted". School-based evaluation systems are to be developed, as are question banks and related research. Although no details are given as to what standards will be developed or followed for the setting up, monitoring and development of these initiatives, it is likely that these will indeed be formulated, and ILTA would do well to remain in contact and correspondence about these developments.

The 1996 Cambridge Syllabus document for English reveals that these examinations are also available in the Caribbean area, Singapore and Brunei, Zambia, Seychelles, amongst others, which means that they and their associated procedures are likely to be widely influential. The document provides the additional information that Cambridge's Council for Examination Development is responsible to the Syndicate (UCLES) "for research and development in all aspects of assessment" and its role is "to ensure that the Syndicate's examinations and tests continue to be valid, reliable and relevant to users generally". No further details are available in this document.

It is evident from documents and correspondence that Mauritius intends to mauritianise its examinations, in "long and fruitful collaboration between the Mauritian Examinations Syndicate and UCLES", and the way in which this proceeds should be of considerable interest to an international language testing association.

==========

NA01: Namibia: "Draft Manual of Standards in English..."

Record:	NA01
TFTSmem:	Alderson
Country:	Namibia
Auth/Pub:	Ministry of Education and Culture
Title(s):	Draft Manual of Standards in English (Second Language) 1993
Language:	English
Contact:	David Forson, Education Officer, European Languages, Ministry of Education and Culture, Private Bag 12026, Ausspanplatz, Windhoek, Namibia. Fax -264.61.222005
Govt/Priv:	govt
Lang/Ed:	lang

StanDef: performance

Comments:

This 60 page document is "a modest attempt to help Grade 10 English teachers to examine and assess the types of questions which they are supposed to train their candidates to communicate effectively and write the examinations well in the future."

The document is an analysis of the 1993 Examination papers in terms of skills, presenting the actual examination papers and mark schemes, samples of pupil performances, with associated marks and comments on the performance, and a copy of the Examiner's Report on Continuous Writing. It contains detailed advice to teachers on how to train their students to perform better. Again, 'standards' is interpreted as standards of performance, but the fact that this sort of manual is issued illustrates a standard of examination practice of relevance to TFTS.

==========

NE01: The Netherlands: "Certificate of Dutch as a Foreign..."

Record: NE01
TFTSmem: Huhta
Country: The Netherlands
Auth/Pub: Nederlandse Taalunie [Dutch Language Union]
Title(s): Certificate of Dutch as a Foreign Language: Objectives, requirements, a short description and underlying principles of the CDFL examinations. (7 pages)
Language: English
Contact:
Govt/Priv: priv
Lang/Ed: lang
StanDef: test

Comments:

Note: This document was compiled from a number of publications and other documents, mainly in Dutch, for comparing the CDFL examination and the Consortium tests. Thus the format of presentation is the same as in the Common Syllabuses at Levels A, B, C and D.

Note: The Certificate of Dutch as a Foreign Language (Certificaat Nederlands als Vreemde Taal) has existed since 1977. Since 1985 it has been under the Nederlandse Taalunie (Dutch Language Union) which is a joint organization between the Netherlands and Belgium.

Objective:
To help to compare the CDFL test with the Consortium tests.

Target group:
Apparently administrators who make decisions about comparabilities of different certificates at schools, universities, workplaces etc.

Procedure:
Unknown. Probably designed at the Dutch Language Union.

Scope of influence:
Probably used at institutions, schools etc. where decisions about the equivalence of the CDFL with other tests have to be made.

Summary:
The document specifies the objectives and content of the test in the same way as the Common Syllabuses at levels A, B, C and D. Included are the following subsections: general and specific objectives of the test, descriptions or lists of the communicative tasks, syntax, morphology and lexis, as well as other linguistics aspects and the test formats.

==========

NE02: The Netherlands/Int'l: [IAEA] [Various material]

Record:	NE02
TFTSmem:	Alderson
Country:	The Netherlands / International
Auth/Pub:	International Association for the Evaluation of Educational Achievement
Title(s):	Various, including The IAEA Guidebook and Methodology and Measurement in International Educational Surveys: The IAEA Technical Handbook ed. John Keeves (1992)
Language:	English
Contact:	Contact: Dr W Loxley, International Association for the Evaluation of Educational Achievement (IAEA) Headquarters c/o SVO, Sweelinckplein 14, 2517 GK The Hague, The Netherlands. Fax: +31 70 360 9951
Govt/Priv:	priv
Lang/Ed:	ed
StanDef:	guideline

Comments:

(See also record NE03.)

This Association has experience that is highly relevant to ILTA's Task Force. The international surveys conducted under their aegis are frequently test-based, and detailed guidelines and standards

appear to exist which govern test construction. As IAEA lists 53 member countries/institutions in its Guidebook, it clearly has considerable experience in dealing with cross-cultural phenomena, presumably including differences in cultural values and 'standards'. Language-related surveys include the recently completed Reading Literacy Study, The Study of Written Composition, and English and French as Foreign Languages surveys, as well as the planned Language Education Survey (see below). We received the following publications:

The IAEA Guidebook (184 pages)

Explains how IAEA works, and gives some details of how standards are monitored. For example, a Technical Advisory Committee monitors the adequacy of technical aspects of all IAEA studies. Pages 33-35 (reprinted from Keeves, see below) give some details of the analysis of test items by content experts, traditional item analysis and Rasch. The document includes a list of the cognitive, attitudinal and perceptual measures used in IAEA studies, and although it does not give details of validity and reliability for such measures, bibliographic references are given where such information can be found.

A brief section (page 46) of the Guidebook draws attention to the difficulties of the construction of measures, and draws the reader's attention specifically to The International Encyclopedia of Educational Evaluation, ed H J Walberg and G D Haertel, Pergamon Press 1990 (in particular articles by R Thorndike, R A Zeller and J B Carroll). In addition, there are numerous references to publications that have resulted from IAEA studies that attest to the standards used in the construction of instruments. A very useful section, extracted from Keeves also, discusses procedures and things to be looked out for, under the heading: "Project Administration, Data Processing and Management Tips for ICCs and NPCs".

The IAEA Language Education Study Proposal, 1993 (76 pages)

This document contains the detailed proposals to conduct an international survey of language education, from 1994 to 1997. It proposes developing "international assessment standards and tests to define basic and fluent levels of communicative competence in major languages, evaluating proficiency in the language (or languages) selected by each participating country" (page 1). In this context, 'standards' appears to mean 'levels of performance'. Later, however, (page 3) the document claims that during the four years of the study, various products will result, including "internationally-validated tests..... along with procedures and common standards for administrating and measuring language proficiency". Later, the claim is repeated in somewhat modified form: The Language Education Study will produce, amongst others, the following product:

> internationally-validated instruments and procedures for

> assessing students' communicative proficiency in English, French, German, and several other languages commonly studied in schools, setting common standards for beginning, communicative proficiency and for advanced, fluent proficiency in these languages along with technical specifications of these instruments for future use (page 12).

The claim is that one result will be "a model framework to guide the development of future language tests throughout the world in various languages" (page 46) as well as "innovations in standard-setting and testing across multiple languages, domains, and language competencies" (page 47). No details are given of how the international validation will be achieved, nor what standards (in ILTA's sense) will be followed, but the following should be noted:

> It is essential that all instruments and procedures are deemed to be valid for each participating country .... Conditions for joining the study require participating countries to adhere to IAEA international codebooks, manuals, sampling instructions, time lines and procedures for test development, test administration, and data processing. (p 69)

Further consultation of these guidelines is needed (see Keeves, below).

The International Coordinating Centre which is responsible for preparation of the tests is the National Foundation for Educational Research in the UK, and the Principle Investigator is Peter Dickson. The institution and the person have good track records for production of tests and research instruments.

Lyle Bachman was involved in early discussions of the Study, and Elana Shohamy and Alister Cumming are on the Steering Committee for the Project, so ILTA should be able to get up-to-date information about existing and developing standards.

In sum, this is an interesting and ambitious project that promises a great deal, and that may result in, and possibly already has, standards we should include.

Following is specific information about The Handbook:

Methodology and Measurement in International Educational Surveys: The IAEA Technical Handbook. ed John Keeves, 1992 (424 pages)

Keeves was Chair of the IAEA Technical Advisory Committee from 1982-89. This Handbook is an editing of existing material "in a way that would report the experiences of IAEA research workers and would provide standards and guidelines, as well as advance appropriate methodology for future IAEA studies" (Foreword, v). Reference. is specifically made to one major source of material for the Handbook being the first edition of the International Encyclopedia of Education, edited by Torsten Husen and T Neville

Postlethwaite. However, the author of the Foreword, T Plomp, Chairman of IAEA, also writes: "The IAEA Technical Handbook should not be seen as the definitive and prescriptive guidebook, but rather as a source of ideas that should inform debate and provide help and support for those who conduct IAEA studies in the future". Nevertheless, in the absence of other documents providing such guidance, it is probably safe to assume that this Technical Handbook does indeed enshrine the practice, procedures and principles that are followed by the Technical Advisory Committee when monitoring IAEA studies and instruments, including language tests. Unfortunately, the Handbook promises more than it delivers, in my view, and could not easily constitute a set of guidelines or standards in its present form. It is therefore of limited immediate relevance to ILTA, but still worthy of consideration because of the nature of the IAEA's work.

The document is far too long to summarise, and ILTA should consult it when/if it decides to draw up its own Standards and Guidelines. A list of its Contents should give some idea of what it contains:

Part I Introduction

1. Survey Studies and Cross-Sectional Research Methods. D A Walker and P M Burnhill, Scotland
2. Longitudinal Research Methods. J P Keeves, Australia
3. Ethical Considerations in Research. W B Dockrell, Scotland

Part II Sampling, Administration and Data Processing

4. Sampling and Administration. M J Rosier and K W Ross, Australia
5. Data Processing and Data Analysis. T N Postlethwaite, Germany

Part III Measurement

6. Scaling Achievement Test Scores. J P Keeves, Australia
7. Reliability. R L Thorndike, USA
8. Validity. J P Keeves, Australia
9. Correction for Guessing. B H Choppin and R M Wolf, USA
10. Item Bias. R J Adams, Australia
11. Attitudes and Their Measurement. L W Anderson, USA
12. Descriptive Scales. E J Eifer, USA

13. Measurement of Social Background. J P Keeves, Australia

14. Observation in Classrooms. L W Anderson, USA

15. Questionnaires. R M Wolf, USA

16. Rating Scales. R M Wolf, USA

Part IV The Analysis of Data

17. Analyzing Qualitative Data. J P Keeves and S Sowden, Australia

18. Missing Data and Non-Response. J P Keeves, Australia

19. Median Polish. E J Kifer, USA

20. Measures of Variation. J P Keeves, Australia

21. Units of Analysis. L Burstein, USA

22. Multilevel Analysis. J P Keeves and N Sellin, Sweden

23. Multivariate Analysis. J P Keeves, Australia

24. Cluster Analysis. D Robin, France

25. Configural Frequency Analysis. R Rittich, Germany

26. Correspondence Analysis. G Henry, Belgium

27. Linear Structural Equation Models. I M E Munck, Sweden

28. Partial Least Squares Path Analysis. N Sellin, Germany

29. Profile Analysis. J P Keeves, Australia

Part V Conclusion

30. Reflections on the Management of IAEA Studies. A Purves, USA

Much of the Handbook is not specific to testing, much less language testing. Indeed, many chapters - especially 6, 7, 8, 10 and all of Part IV - read like standard textbooks. Their value is that they are published in a Technical Handbook by an organisation that constructs tests, and constitute some kind of guidelines for that organisation, although they are certainly not couched in the form of checklists.

Some chapters more than others refer to practice and procedures within IAEA: chapters 9, 12, 14 and 15 in particular. With respect to tests specifically, of most relevance are pages 86-90

in chapter 4 on the administration of a testing program, and page 99 in chapter 5 on test data. Dockrell's chapter on Ethics makes explicit reference to the AEA/APA/NCME stuff, and the final chapter on Management contains reflections on cross-cultural issues, the management of studies, and the relationship among projects, that may have relevance to an international organisation like ILTA. This final chapter contains a number of recommendations, which will not be repeated here, but which may well be of general interest to an international organisation like ILTA, though not of direct relevance to the Task Force (pages 420-424).

==========

NE03: The Netherlands/Int'l: [IAEA] "Standards for Design..."

Record: NE03
TFTSmem: Huhta
Country: The Netherlands / International
Auth/Pub: International Association for the Evaluation of Educational Achievement
Title(s): Standards for the Design and Operations in IAEA Studies. Prepared by Andreas Schleicher, Director of data management and analysis, IAEA Secretariat. Publisher: Statistics Sweden 1994. (23 pages)
Language: English
Contact:
Govt/Priv: priv
Lang/Ed: ed
StanDef: guideline

Comments:

(See also record NE02.)

Objective:
Argues for the need to establish standards for the design and data collection in the IAEA studies. Outlines the aspects in the studies which need to be standardized and offers some suggestions as to the content of such standards. The aim is thus to increase the comparability between different studies and the measurement of educational achievement over time. The idea is that each study would then specify its study-specific standards on the basis of these (more general) standards.

Target group:
Those planning and carrying out IAEA studies.

Procedure:
Prepared by the Director of data management and analysis of the IAEA. Based on experience from several previous IAEA studies, because the writer often cites examples from these studies to illustrate the nature of problems caused by too vague or lacking

standards. (perhaps there has been a seminar on this issue as the subtitle of the document suggests)

Scope of influence:
This document is an argument (and suggestions) for the creation of some general standards for the IAEA. As such it appears to be targeted to the IAEA 'community'.

Summary:
The following is the original summary of the document:

"This paper focuses on technical standards that are related to the design and data collection operations of IAEA studies. It does not cover standards in the more substantive aspects of IAEA studies, such as the development of operationalized research questions and the construction of the data collection instruments."

The document first argues that the changing needs require permanent standards for the IAEA studies (results from different studies need to be linked, ed. achievement needs to be measured over time). Also, the credibility of the studies requires that minimum standards are specified and applied. Standards can ensure e.g. that
- populations and entities of reporting are in fact comparable
- definitions of variables and their operationalizations are valid and equivalent
- measurement of variables is consistent over time and across countries
- data collection and management are comparable
- data analyses match the type and quality of data and the reporting requirements
- various survey constraints have a similar impact in different systems

The document presents the following issues, discusses related problems and offers lists of points that the standards should cover:

Standards and Target Populations
- internationally desired populations, nationally defined, excluded and achieved populations (should all be clearly defined and reported)
- coverage of educational sub-systems in countries
- population exclusions: of schools, of students
- target grades and target ages
- age-based and grade-based target populations
- single-grade versus multi-grade target populations

Standards and the Units of Sampling, Analysis, and Reporting
- defining the units of sampling
- sub-sampling of classes and sub-sampling of students
- definitions of a 'country'

Standards for Accuracy and Precision

Standards for Response Rates for Schools

Standards for Data Collection Operations
- relation between data collection operations and study design

Standards for the Data Management

Standards for Data Analysis
- standards for handling of missing data

==========

NZ01: New Zealand: "Regulations and Prescriptions Handbook..."

Record:	NZ01
TFTSmem:	Alderson
Country:	New Zealand
Auth/Pub:	New Zealand Qualifications Authority
Title(s):	Regulations and Prescriptions Handbook and Appendix to Examiner's Contract.
Language:	English
Contact:	Peter Morrow, Assessment and Moderation Officer, New Zealand Qualifications Authority, U-Bix Centre, 79 Taranaki Street, Wellington, New Zealand. P O Box 160 Fax (04) 802 3112
Govt/Priv:	govt
Lang/Ed:	ed
StanDef:	guideline

Comments:

Essentially specifications for examinations used by examiners to prepare written examinations in forms 5-7, and also used by teachers for school based assessments which form part of the national awards. Includes details of all examinations, and NZQA Regulations and Timetables. Under Regulations can be found useful details of the Administration of Assessment Procedures, relating to Setting Examinations, External Examination Centres, Examination Supervision, Marking Examination Answer Booklets, Breaches of the Rules and so on.

Appendix to Examiner's Contract. Summarises the requirements to which examiners must adhere when setting a paper.

Materials provided by the Examinations Coordinator for the guidance of examiners, said to be "the beginning of an on-going process of standardisation and development":

i) Style Booklet. Style, Format and Design of Examination Papers. Sections on Fonts; Text; Illustration; Proofreading; reasons for various formats; Multiple-choice questions; Question and Answer

Booklet Format; Mark Transfer System

ii) Trade Certificate and AVA Examinations 1994: Guidelines for Examiners and Moderators. Includes sections on: Introduction (general objectives, overview of the process, timelines for submission and production, responsibilities of examiners and moderators, confidentiality and security, submission of papers, fees); Setting an Examination Paper; Types of Questions that May be Used in an Examination; The Marking schedule; Comments from the Editors; Moderation; The Materials List; Proof-reading; sample Checklists and Report Forms

A very useful and thorough document, partially addressing our concerns.

==========

PO01: Portugal: [Materials to support teachers nationwide...]

Record: PO01
TFTSmem: Alderson
Country: Portugal
Auth/Pub: Institute for Educational Innovation, Ministry of Education
Title(s): Materials (in Portuguese) to support teachers nationwide in "developing assessment materials and procedures consistent with the new assessment system in Portugal". Includes: 30+ page booklet Avaliar e Aprender, explaining the new evaluation system, and the reasons for change.
Language: Portuguese
Contact: Dr Domingos Fernandez, Instituto de Inovacao, Ministerio Da Educacao, Travessa Terras de Sant'Ana No 15, 1200 Lisboa, Portugal
Govt/Priv: govt
Lang/Ed: ed
StanDef: other

Comments:

Training materials on four main topics: Perspectives on Evaluation, Formative Evaluation, Summative Evaluation, Differentiated Pedagogy and Educational Support.

It is unclear what force the principles enunciated in the booklets have in practice, nor how in detail they relate to assessment practice and procedures being developed in Portugal.

==========

SA01: South Africa: "Handbook for English: GEC Exam..."

Record: SA01

TFTSmem: Turner
Country: South Africa
Auth/Pub: Independent Examination Board, Johannesburg
Title(s): Handbook for English: GEC Examination. (1994)
Language: English
Contact: David Adler, National Director, The International Examinations Board, 24 Wellington Road, Parktown 2193, South Africa. Tel: (011) 643-7098
Govt/Priv: priv
Lang/Ed: lang
StanDef: test

Comments:

Explanatory Note regarding records SA01 and SA02:

Established in 1988, the IEB is an independent, non-profit organisation which provides examinations for schools and adult education programs. The recent developments in South Africa have led to the formulation of an integrated national policy framework for education and training. The IE3 is involved in syllabus design and assessment. The new framework will have a General Education Certificate (GEC) for both regular school age and adult students, at the end of the period of free compulsory education; Further Education Certificate (FEC) for both regular and adult students, at the tertiary entrance level; and Higher Education Certificates.

Objective/Purpose:
To provide guidance in the areas of curriculum and assessment during the transition period toward a national policy framework. To assist in content based on program objectives that will serve as the core of the General Education Certificate Examination in 1994.

Target group/Intended audience:
All educators across South Africa in both regular (formal) schools and adult section schools.

Procedure:
Not specified. It was stated, however, that the IEB aims to modify the traditional secrecy about assessment criteria, marking and moderating procedures and to involve teachers. Practicing teachers were asked to put forward their names to be involved.

Scope of influence:
Not easy to know from document, but supposedly same as 'Target group/Intended audience' above.

Summary:
The document is divided into two parts: English first language and English second language. The IEB states that language is a highly contested area in education in South Africa and it dislikes the terms first and second language. Because these are

currently approved terms, however, and generally understood the IEB has decided to employ them at present. The document divides each part into five sections: introduction, syllabus, guidelines for assessment, examination format, and exemplars of examination tasks. Precise examination format and times are provided along with several specific task examples. It is stated that these adhere very closely to the curriculum (syllabus). Teachers are encouraged to follow this pattern in classroom activities and assessment.

==========

SA02: South Africa: "User Guide 1, General Handbook: Adult..."

Record: SA02
TFTSmem: Turner
Country: South Africa
Auth/Pub: Independent Examination Board, Johannesburg
Title(s): User Guide 1, General Handbook: Adult Basic Education (ABE) Level 3 (1994).
Language: English
Contact: David Adler, National Director, The International Examinations Board, 24 Wellington Road, Parktown 2193, South Africa. Tel: (011) 643-7098
Govt/Priv: priv
Lang/Ed: ed
StanDef: performance

Comments:

(See also 'Explanatory Notes regarding records SA01 and SA02' in the 'Comments' field of record SA01.)

Objective/Purpose:
To explain the Pilot Examination project which has been set up to answer two questions: (1) What appropriate and achievable outcomes can and should be expected at the various stages of ABE in South Africa? and (2) What standards should be demanded at these different stages, standards that are in the reach of adult learners and ABE providers, but that also encourage and assist in the development of purposeful, quality learning?

Target group/Intended audience:
All agencies, organisations and adult education centers who will be entering candidates in the examination.

Procedure:
The document and the Pilot Examination project were developed by the IEB. The process included consultations with interest groups, employer and labor groups from industry, people from the academic world, those involved in national policy debates and the Non-Government Organisation sector. A steering committee and working groups for examination content were set up. These groups

included members drawn from the various sectors in adult education.

Summary:
The document is the first USER GUIDE in a series of four which will introduce and guide educators through the examination process of a specific subject domain. This particular USER GUIDE discusses The Pilot Examination project which covers two subject domains for ABE: Communications in English and Mathematics.

==========

SA03: South Africa: "Standards -- The Loaded Term..."

Record: SA03
TFTSmem: Turner
Country: South Africa
Auth/Pub: Mamphela Ramphele, Deputy Vice-Chancellor of University of Cape Town.
Title(s): Standards -- The Loaded Term: Occasional papers, No. 1
Language: English
Contact: Nan Yeld, Academic Support Programme, University of Cape Town, Rondebosch, 7700, South Africa. E-mail Yeld@socsci.uct.ac.za
Govt/Priv: priv
Lang/Ed: ed
StanDef: other

Comments:

Summary:
This paper addresses the rising concern about standards of performance, particularly in the academic context in South Africa at such institutions as the University of Cape Town (UCT). It discusses the following topics: the need to draw upon lessons from history in relation to the standards debate; current standards at UCT; and suggestions for future actions towards a commitment to equal opportunity and quality education. Suggestions are: to continue setting both entrance and exit standards appropriate to each discipline, and to articulate them clearly; to encourage a culture of excellence in performance; and, to establish and communicate unambiguously measures of excellence in standards.

==========

SD01: Sweden: "[Central examinations in the senior ...]"

Record: SD01
TFTSmem: Huhta
Country: Sweden
Auth/Pub: Department of Education and Educational Research;

Gothenburg University

Title(s): a) Centrala prov in gymnasieskolan 1993/94 - Allmänna instruktioner ('Centrala prov' ('central examinations') in the senior secondary schools 1993/94 - General instructions)

b) Instruktionshäfte för centrala prov in moderna språk i åk 2 på gymansieskolans treåriga linjer (Instructions for the 'centrala prov' in modern languages for the second grade of the 3-year senior secondary school). Spring 1994. (12 pages)

c) Centralt prov in engelska. Årkurs 2:3 1994. Mall och bedömingsinstruktioner (The 'central prov' in English. Grade 2:3, 1994. Model answers and assessment instructions.) (8 pages)

d) Bedömning av delprovet Uppsats i gymnasieskolans åk 2:3. (Assessment instructions for the subtest 'Essay writing' in grade 2:3 of the senior secondary school). Spring 1994. (8 pages)

e) Normer och justeringstabell för centralt prov i engelska åk 2:3, 1994 (Tables for norming and adjusting (the certificate grades) for the 'centralt prov' in English, grade 2:3, 1994) (4 pages)

Language: Swedish

Contact: Department of Education and Educational Research; Gothenburg University; Box 1010; S-431 26 Moelndal; Sweden

Govt/Priv: govt

Lang/Ed: lang

StanDef: test

Comments:

'Centrala prov' are centrally organised examinations in several subject matters, including the foreign languages (English, German, and French), for the senior secondary schools in Sweden (students are about 16 to 18 years old). The exams take place at the end of the second year both in the more vocationally oriented 2-year schools and the more academically oriented 3-year schools. Examinations are designed by a special group of teachers and test designers at the Gothenburg University. The National Board of Education and the Gothenburg U. are jointly responsible for the exam. Teachers do the marking, send a certain percentage of the papers and results to the Univ. of Gothenburg where the test is normed; the University then sends guidelines back to the teachers on how to adjust the grades they give to their students.

There is also a fairly similar testing system for the Swedish junior secondary school called 'standardprov'. Unlike the 'central prov', this is not an obligatory test.

There are no written guidelines for test construction, apparently because the group who designs the tests is not very large and they meet regularly to review each other's work.

The norm-referenced 'standard' and 'centrala prov' testing systems will be replaced by a new criterion- referenced (objectives or proficiency referenced) system in 1995/96.

Objectives and target groups:
To help classroom teachers to mark the 'centrala prov' subtests and to adjust their grades according to the norm-related information from the test designers.

Procedure:
These documents are produced for each administration of the examination. They are apparently produced by the group of test designers and analysers at Gothenburg University.

Scope of influence:
The 'centrala prov' and the guidelines accompanying each examination are designed by the National Board of Education (Skolverket) together with Gothenburg University. Thus they have a considerable influence within the Swedish school system and the teachers working there.

Summary:
a) Centrala prov in gymnasieskolan 1993/94 - Allmänna instruktioner

The document deals with general information on how the examination is to be administered at schools and how the norming and adjustment of the certificate grades is done on the basis of information from the examination.

b) Instruktionshaefte ...

The document contains guidelines on how to administer the three foreign language tests of the 'centrala prov' system (English, French, German). Instructions are given on the various practical considerations the test administrations require: e.g. seating of the students, instructions to the students before and during the tests, breaks between tests. The process of marking and sending a selection of test papers to be normed is also described. In addition, the document contains more general information about matters such as secrecy of the test material and archiving some of the material at school

c)... Mall och bedömningsinstruktioner (engelska)

The document contains the right answers to the multiple-choice and gap-filling tests that cover reading and listening comprehension and vocabulary and grammar. For the gap-filling tests, lists of 'right', 'acceptable', and 'wrong' answers are given. There

are similar documents for the other languages.

d) Bedömning av delprovet Uppsats ...

The document includes a brief description of the three grade-levels and the assessment criteria. Most of the pages contain examples of essays representing the different grades with short commentaries about the reasons for the grades awarded. (Separate versions exist for the three languages tested.)

e) Normer och justeringstabell ...

The document contains a blank table for adjusting the certificate grades. It also includes descriptive information about the test results (based on the subpopulation of the test takers with which the norming is carried out; in this case about 3,600 students in this case), such as the means and standard deviations broken down by the subtest and the sex, and a table of subtest correlations. (Separate versions of this document exist for the three languages.)

==========

SE01: Seychelles: "The Certificate of Proficiency in Engl..."

Record:	SE01
TFTSmem:	Alderson
Country:	Seychelles
Auth/Pub:	Ministry of Education and Culture
Title(s):	The Certificate of Proficiency in English as a Second Language
Language:	English
Contact:	G Vidot, Research and Evaluation Section, Ministry of Education, P O Box 48, Victoria Mahe, Seychelles. Fax 224859.
Govt/Priv:	govt
Lang/Ed:	lang
StanDef:	test

Comments:

Describes a newly developed Certificate Programme used at the Seychelles Polytechnic, with details of weighting, continuous assessment, aims and objectives, skills to be mastered and proficiency levels: guidelines for assessment of student work, for listening, reading, speaking and writing, divided into 7 levels and a number of 'competency areas' (criteria). A helpful covering letter explains the work of the Curriculum Development section of the Ministry and its involvement in assessment and testing. It points out that the "rules, procedures and standards for the examinations are largely unwritten..(but that) it is becoming very important to establish and formalise rules, procedures and standards for

the examinations. To date, chief examiners are being guided by past papers in their subject. The only document I have been able to trace which may at least partly answer to the description in your letter" is this document.

==========

SI01: Singapore: "PSLE Information..., Assessment Guide..."

Record: SI01
TFTSmem: Alderson
Country: Singapore
Auth/Pub: Examinations and Assessment Branch, Ministry of Education
Title(s): i) PSLE Information Booklet: English Language 1994
ii) Assessment Guidelines (Lower Secondary English Language, Sec 1 and Sec 2) 1994
Language: English
Contact: C Seng, Ministry of Education, P O Box 746, Kay Siang Road, Singapore 1024, Republic of Singapore. Fax: 4792878
Govt/Priv: govt
Lang/Ed: lang
StanDef: test

Comments:

These documents are confidential, for the use of TFTS only. Document i) describes English language papers for the Primary School Leaving Examination, giving details of purpose, format, duration, the specifications and sample questions and marking scheme.

Document ii), despite having a different title, is essentially the same, but for lower secondary school pupils. It is divided into: objective and duration, format, specifications, marking scheme and sample questions.

==========

SW01: Switzerland: "Language B Guide, First Edition, 1994."

Record: SW01
TFTSmem: Alderson
Country: Switzerland
Auth/Pub: International Baccalaureate
Title(s): Language B Guide, First Edition. 1994
Language: English
Contact: International Baccalaureate Organisation, Route des Morillons 15, 1218 Grand Saconnex, Geneva, Switzerland
Govt/Priv: priv

Lang/Ed: lang
StanDef: other

Comments:

A description of the Language B Programme to be introduced as part of the IB from May 1996. Includes Aims, Objectives, Syllabus Outline, Syllabus Guidelines, Assessment Outline, and Assessment Details and Criteria. Includes descriptions of possible exam formats, and the criteria used to assess work and to define levels of performance. Of interest only because it is international and influential, and it may be that the IB plan to produce a document setting out codes of practice in due course.

==========

SW02: Switzerland: "Pedagogic policy..., Guidelines for..."

Record: SW02
TFTSmem: Huhta
Country: Switzerland
Auth/Pub: Eurocentres
Title(s): a) Pedagogic Policy Statements: Pedagogic Policy 6. Levels, Assessment, Certification (9 pages); b) Guidelines for Teachers (1993) (3 pages)
Language: English
Contact: Eurocentres; Head Office; Seestrasse 247; CH - 8038 Zurich; Switzerland; tel. 01 485 52 00; fax. 01 481 61 24
Govt/Priv: priv
Lang/Ed: ed
StanDef: performance

Comments:

Objectives and target groups:
Both documents: to help foreign language teachers in the Eurocentre schools to follow the Eurocentres' "system for defining levels, for assessing student level and progress, and for certifying achievement"

Procedure:
The documents are produced by the Eurocentres foundation.

Scope of influence:
Apparently the schools within the Eurocentres system.

Summary:
a) Pedagogic Policy 6. Levels, Assessment, Certification

The document presents "ground rules for definition of levels, assessment and certification in Eurocentres courses". First, the

Eurocentres' 10-level system is explained, including estimates on the amount of instruction typically needed to proceed from one level to the next. Second, the assessment and feedback procedures are specified: types of activities or tests used for placement, diagnostic (in-course), and certification purposes. These guidelines vary depending on the nature of the language courses (long vs. short courses, intensive or holiday courses). The role of the Eurocentres' tests and assessment procedures is explained (the Itembanker, RADIO, and LOC, see below).

There are separate sample certificates, video samples for oral assessment, and four background papers that are referred to in this document, but are not part of it. The four background papers cover 1) assessment of oral proficiency ("RADIO": Range, Accuracy, Delivery, Interaction), 2) testing of system knowledge ("ITEMBANKER", an IRT-based test production programme with a bank of 1000 items covering vocabulary, grammar and cohesion), 3) assessment of written tasks ("LOC": Language, Organisation, Communication), and 4) a tutorial/personalised record card system ("feuille de route").

b) Guidelines for Teachers:
The document explains in more detail and at a more practical level the assessment procedures and activities referred to in document a). It contains advice on how to give feedback to students and explains the assessment of speaking and writing skills, as well as the Itembanker programme. The section on assessing oral proficiency (RADIO) includes advice on the task types and on the assessment procedure. The section on assessing writing (LOC) presents a model of how to quickly assess a whole class of students.

==========

TA01: Tanzania: "Continuous Assessment: Guidelines..."

Record:	TA01
TFTSmem:	Alderson
Country:	Tanzania
Auth/Pub:	National Examinations Council of Tanzania
Title(s):	Continuous Assessment: Guidelines on the conduct and administration of continuous assessment in secondary schools and teacher training colleges. 1989
Language:	English
Contact:	P P Gandye, Executive Secretary, The National Examinations Council of Tanzania, P O Box 2624, Dar es Salaam, Tanzania
Govt/Priv:	govt
Lang/Ed:	ed
StanDef:	guideline

Comments:

The purpose of the guidelines is to "redress the weaknesses found in the conduct and administration of continuous assessment in schools." It contains guidelines on test preparation, scoring and recording of test scores, project work assessment (including how to timetable, assess and give feedback), and character assessment (an important part of Tanzanian education).

==========

UG01: Uganda: "Language Examinations, Ordinary and Adv..."

Record: UG01
TFTSmem: Alderson
Country: Uganda
Auth/Pub: Uganda National Examinations Board
Title(s): Language Examinations, Ordinary and Advanced level, English, Luganda, Arabic
Language: English
Contact: Dr C I Cele, Uganda National Examinations Board, P O Box 7066, Kampala, Uganda
Govt/Priv: govt
Lang/Ed: lang
StanDef: performance

Comments:

The document is a copy of the syllabuses for these languages. However, in an accompanying letter, Dr Cele writes: "We have detailed procedures on testing in general, but this would take me a while to compile. Currently a draft on 'STANDARDS FOR EXAMINATIONS' has been drafted and is awaiting discussions to see how it relates with the current procedures the Board uses." We suspect this is fairly typical, in that internal rules, procedures and standards seem to exist, but are not written down, or drawn up into codes of practice, or similar.

==========

UK01: England: "Handbook for Centres: All Cambridge Exams..."

Record: UK01
TFTSmem: Huhta
Country: England
Auth/Pub: University of Cambridge Local Examinations Syndicate (UCLES)
Title(s): Handbook for Centres. All Cambridge Examinations (International, revised 1994)
Language:
Contact: The Publication Office, UCLES, 1 Hills Road, Cambridge CB1 2EU, UK

Govt/Priv: priv
Lang/Ed: ed
StanDef: guideline

Comments:

The Handbook is a very thorough and detailed guide on practical test administration including the preparations and paperwork needed. The document hardly misses any factor that might affect the way the examination is conducted and the candidates treated. Evidently a document based on a long experience on test administration.

Objectives and target groups:
For those responsible for administering Cambridge examinations (EFL, as well as other examinations). To ensure standardised conditions for test administration and for the treatment of test takers.

Procedure:
Not specified (but produced by UCLES)

Scope of influence:
Test centres for Cambridge exams all over the world.

Summary:
The document has the following subsections (main content briefly explained):

1. General information
2. Method of entry
   Specifies e.g. entry documents, fees, transfers of candidates, late entries, withdrawals.
3. Arrangements for the examination
   Specifies what material the centre will receive from UCLES and what stationery the centre must provide for the candidates.
4. Instructions for the conduct of examination
   Explains in detail what the centre must do prior to, at the beginning, during, at the end, and after, the examination. These include instructions on e.g. safe custody of question papers, seating, invigilation, identification of candidates, instruction and advice to the candidates, late arrivals, irregularities, emergencies, packing of scripts.
5. Issue of results
   E.g. on provisional results, duplicate copies of certificates, Data Protection Act.
6. Regulations governing provisions for candidates who are handicapped or affected by adverse circumstances.
   Gives a very detailed descriptions of what to do in these situations; e.g. lists of acceptable and unacceptable reasons that are counted as 'temporary adverse circumstances' that may be given special consideration

when marking the papers.

The document also contains appendices e.g. on the use computers / word processors, use of a reader, an amanuensis, or a practical assistant (for some disabled candidates), and checking the listening comprehension tapes. One of the appendices describes how the above matters are similar or different as far as EFL examinations are concerned.

==========

UK02: England: "The Common Syllabuses at Levels A, B, ..."

Record: UK02
TFTSmem: Huhta
Country: England
Auth/Pub: Consortium for the European Certificate of Attainment in Modern Languages
Title(s): The Common Syllabuses at Levels A, B, C and D. (in different languages: English, French, German, Greek, Italian, Spanish) (14 pages)
Language: (see 'Title(s)' above)
Contact: Secretariat
Consortium for the European Certificate of Attainment in Modern Languages
University of London Examinations and Assessment Council
32 Russell Square
London WC1B 5DN
UK
Govt/Priv: priv
Lang/Ed: lang
StanDef: performance

Comments:

Note: the Consortium is "a formally constituted partnership of prestigious institutions with a common interest: the teaching and testing of languages for non-native speakers". Their work began in 1989 first under the ERASMUS programme, then under the LINGUA programme of the European Community. The aim is to include all official languages of the EC. The consortium has agreed to a) promote the teaching, learning and testing of EC languages both within and beyond Europe, b) promote this objective through developing tests and awarding certificates to successful candidates, and c) establish equivalencies (see point 2 'Objective' below for details).

The member institutions:

Aristotle University of Thessaloniki
Centre International d'Études Pédagogiques de Sèvres
Fachverband Deutsch als Fremdsprache

Universidad de Granada Centro de Lenguas Modernas
Università per Stranieri de Siena
University of London Examinations and Assessment Council

Objective:
One of the aims of the Consortium is "To help to establish equivalence between our common syllabus, its associated tests and awards and other language tests offered by our member institutions. The common syllabus has common assessment objectives and common test formats. Where tests already exist in certain languages, these tests are being granted equivalence on the basis of their assessing the same or comparable objectives."

Target group:
Not specified in the document, but apparently all involved in designing tests and making administrative decisions about e.g. test equivalences.

Procedure:
Not yet known to me. Apparently a joint committee of the member institutions has designed the common syllabus.

Scope of influence:
The EC countries concerned?

Summary:
The Common Syllabuses at Levels A, B, C and D presents the following information and descriptions for the four proficiency levels:
A. General objectives: highlight the characteristics of the language to be used and tested at each level
B. Specific objectives: objectives for the four skills are presented (reading, writing, speaking, listening)
C. Communicative tasks: defined and listed, with examples
D. Syntax, morphology and lexis: brief descriptions of the kinds needed
E. Other linguistic aspects: e.g. regional varieties, speed of delivery
F. Test format: types and lengths of tasks, mode of delivery and weighting for each of the four skills are described

==========

UK03: England: "Issues in Public Examinations (1991)"

Record: UK03
TFTSmem: Alderson
Country: England
Auth/Pub: Luijten (Ed)
Title(s): Issues in Public Examinations (1991)
Language: English
Contact: The publisher, Uitgeverij Lemma B V, Postbus 3320, 3502 GH Utrecht, The Netherlands

Govt/Priv: priv
Lang/Ed: ed
StanDef: other

Comments:

A selection of the proceedings of the 1990 IAEA Conference, which "reflects the attempts by examining bodies and institutes to improve examination systems and examining techniques, to develop reliable instruments and to establish standards in public examinations".

==========

UK04: England: "Examinations: Comparative and International..."

Record: UK04
TFTSmem: Alderson
Country: England
Auth/Pub: Eckstein and Noah (Eds.)
Title(s): Examinations: Comparative and International Studies. (1992)
Language: English
Contact: The Publisher: Pergamanon Press plc, Headington Hill Hall, Oxford OX3 0BW, UK
Govt/Priv: priv
Lang/Ed: ed
StanDef: other

Comments:

Several chapters comparing national systems of examinations; examination systems in Africa: between internationalization and indigenization; a comparison of Mediterranean and Anglo-Saxon Countries in tradition and change in national examination systems.

==========

UK05: England: "GCSE Mandatory Code of Practice..."

Record: UK05
TFTSmem: Alderson
Country: England
Auth/Pub: Schools Examination and Assessment Council
Title(s): GCSE Mandatory Code of Practice, January 1993
Language: English
Contact: Colin Robinson, Head of Evaluation and Monitoring Unit, Schools Curriculum and Assessment Authority, Newcombe House, 45 Notting Hill Gate, London W11 3JB, England Fax: 0171 221 2141
Govt/Priv: govt
Lang/Ed: ed

StanDef: guideline

Comments:

Reviewed in Alderson, Clapham and Wall. (forthcoming) Language Test Construction and Evaluation. Cambridge UP.

This entry needs considerable supplementation by reference to Alderson et al. as this is a significant document.

Note that the original 1993 document has been slightly revised in March 1994 by introducing "rules to govern the award of grade A and improvements in the way assessment of spelling, punctuation and grammar is dealt with by question paper examiners and coursework moderators. Appendices dealing with candidate malpractice and timetable clashes have also been added. "Otherwise the document is the same as the 1993 version.

==========

UK06: England: [various material from Univ. London]

Record: UK06
TFTSmem: Alderson
Country: England
Auth/Pub: University of London School Examinations Board
Title(s): [various]
Language: English
Contact: Anne Rickwood, Graded Test Development Officer, University of London School Examinations Board, Stewart House, 32 Russell Square London, WC1B 5DN England Fax 0171 631 3369
Govt/Priv: priv
Lang/Ed: lang
StanDef: test

Comments:

Information, including specimen test, on Certificates of Attainment in English, for non-native speakers. Covering letter claims that "such tests testify more accurately to the actual ability of the intending undergraduate to use the language than do the more widely known multiple-choice or sentence-completion type of tests. We hope to be able to prove this once the tests have been in existence long enough for students who have taken them to complete their undergraduate education. "Reference is also made to planned research to "prove comparability between performance on our Certificate of Attainment and on TOEFL" (which rather contradicts the previous assertion!). No references to standards or validation procedures are made.

==========

UK07: England: [LCCI] "Centre Application Form & Fee Sheet"

Record: UK07
TFTSmem: Alderson
Country: England
Auth/Pub: London Chamber of Commerce and Industry (LCCI)
Title(s): Centre Application Form
Fee Sheet
Guide to LCCI Qualifications
List of LCCI NVQ
Details of the Award
Guide for Centres
Language:
Contact: Barnaby Elphick, NVQ Manager, London Chamber of Commerce and Industry, Marlowe House, Station Road, Sidcup, Kent DA15 7BJ England. Fax 0181 302 4169/ 0181 309 5169
Govt/Priv: priv
Lang/Ed: ed
StanDef: guideline

Comments:

Sent glossy publicity materials, in the form of an "information pack", including material given under 'Title(s)' above. The Guide for Centres contains some details of quality assurance procedures for the LCCI Examinations Board's National Vocational Qualifications (NVQs).

It describes the NVQ Framework (which is a national framework) and the Quality Assurance Model, which essentially consists of 2 verification visits per Centre and per Award (examination) per year by regional and local External Verifiers. The reports of such visits are monitored by LCCI NVQ Managers, and made available to Centres. The Guide contains some information on the verification process, and the certified training up to national standards required for assessors, advisers and verifiers. It details the responsibilities for Centres, the Centre approval process, and the main points of the nationally agreed NCVQ Common Accord for Awarding Bodies, which was introduced to "ensure standardisation of procedures and principles across NVQ/GNVQ Awarding Bodies". These relate to Management Systems, Administrative Arrangements, Physical Resources, Staff Resources, Assessment, Quality Assurance and Control and Equal Opportunities and Access policies. The document also gives guidelines on Special Needs, the LCCI's Equal Opportunities Policy, and their implementation strategies and appeals procedures.

As an Appendix, it includes Guidelines on Assessment of National Vocational Qualifications, covering the process of evidence collection, presentation, assessment and verification for NVQs, which draw on a National Guide dated 1991. Topic headings include: Access to Assessment, Assessment Methods, Evidence Criter-

ia (which include 'sufficiency, validity, authenticity and currency'), Collection and Presentation of Evidence, Portfolio, and Key Roles in Assessment and Verification.

==========

UK08: England: "Introduction to the National Language Stds..."

Record: UK08
TFTSmem: Alderson
Country: England
Auth/Pub: Languages Lead Body
Title(s): Introduction to the National Language Standards (May 1993); National Language Standards: Breaking the Language Barrier Across the World of Work (May 1993)
Language: English
Contact: Languages Lead Body, C/O CILT, 20 Bedfordbury, London WC2N 4LB, England. Fax: 0171 379 5082
Govt/Priv: priv
Lang/Ed: lang
StanDef: performance

Comments:

Aim is to set standards for language of performance, to improve 'the economic performance of this country': "The standards will guarantee employers and those learning a language that courses and training are precisely what business requires"

Describes a framework for language qualifications, within the larger framework of standards for vocational qualifications "from construction to administration, from management to hairdressing". The standards are intended to apply to language skills and training in any language, including EFL. They are set at five levels, in the four macro skills of listening, speaking, reading and writing, and are intended to guide test specifications. They are divided into : elements "which describe what someone can achieve using a foreign language"; performance criteria "which indicate what has to be demonstrated to show competence"; assessment guidance; and range statements "which define the instances in which evidence of competence is required"

They are 'recognised' by the National Council for Vocational Qualifications and the Scottish Vocational Educational Council, and beginning to be used by examinations bodies like City and Guilds, London Chamber of Commerce and Industry, RSA and so on as well as by employers.

There is no reference in the documents to procedures for ensuring that the qualifications based on these 'standards' are valid or reliable, although the word 'validation' does appear in

the glossary, as follows: "A measure of the extent to which learning objectives relate to the needs they purport to address, and the extent to which learning that occurs is related to the agreed learning objectives. In assessment, a measure of the extent to which an assessment assesses what it purports to assess". Unfortunately the documents contain no guidance on how validation will, should be conducted on these 'standards'.

==========

UK09: Wales: "Syllabuses for French, German, and Span..."

Record: UK09
TFTSmem: Alderson
Country: Wales
Auth/Pub: Welsh Joint Examination Committee, Cardiff
Title(s): Syllabuses for French, German and Spanish at A, AS and GCSE levels.
Language: English
Contact: Derec Stockley, Assistant Secretary, Welsh Joint Education Committee, 245 Western Avenue, Cardiff CF5 2YX, Wales
Govt/Priv: govt
Lang/Ed: lang
StanDef: performance

Comments:

These documents contain nothing on standards, simply detailing the aims, assessment objectives, weighting, and content (in terms of topics, set books, functions, notions, grammatical structures and vocabulary).

==========

UK10: England: "The BPS Statement and Certificate..."

Record: UK10
TFTSmem: Alderson
Country: England
Auth/Pub: The British Psychological Society
Title(s): The BPS Statement and Certificate of Competences in Occupational Testing
Language: English
Contact: Colin Newman, Executive Secretary, The British Psychological Society, St Andrews House, 48, Princess Road East, Leicester LE1 7DR, England. Fax: 01533 470787. e-mail: BPS1@le.ac.uk
Govt/Priv: priv
Lang/Ed: ed
StanDef: guideline

Comments:

The Society has a Steering Committee on Test Standards which has published a number of guidance statements for the Society, listed below, and has been proactive, within occupational testing, in establishing a Certificate of Competence in Occupational Testing. The Executive Secretary of the BPS writes: "We have every intention of extending the scheme to include educational testing in the next few years, as soon as the relevant work can be done by our volunteer committee. We worked on occupational testing initially as it is in that field that it seemed to us there were the greatest dangers of standards not being maintained in the UK."

The BPS Statement and Certificate of Competences in Occupational Testing, available since 1991, is awarded to those who have had their competencies checked in seven main areas:

defining assessment needs
basic principles of scaling and standardisation
reliability and validity
deciding when tests should be used
administering and scoring tests
making appropriate use of test results
maintaining security and confidentiality

These are subdivided into 97 elements of competence, all of which must be passed, according to a 'properly qualified Chartered Psychologist'.

We hold the Information Pack about this Certificate. It is a very informative and useful document, which is rather hard to summarise.

The Steering Committee on Test Standards publish Psychological Testing: A Guide. This 17-page document offers guidance for non-psychologist test users in educational, clinical and occupational fields. It has three sections: an Introduction to testing, different applications and various quality control issues; Practical advice on what to look for in a test; Further information: where to go from here (which includes a short bibliography). The first section is a very brief textbook on testing, but the second section "amounts to a guide as to what to expect from a good test manual", and comes close (in four pages) to what ILTA is interested in identifying.

In addition, there is a division of the BPS entitled Division of Occupational Psychology, which publishes a Code of Professional Conduct, freely available to members of the general public

The BPS also issue a "non-evaluative list of publishers and distributors of psychological tests in the UK.

==========

UK11: England: [BPS] "Psychological Testing: A Guide"

Record: UK11
TFTSmem: Alderson
Country: England
Auth/Pub: The British Psychological Society
Title(s): Psychological Testing: A Guide
Language: English
Contact: Colin Newman, Executive Secretary, The British Psychological Society, St Andrews House, 48, Princess Road East, Leicester LE1 7DR, England. Fax: 01533 470787. e-mail: BPS1@le.ac.uk
Govt/Priv: priv
Lang/Ed: ed
StanDef: guideline

Comments:

The Steering Committee on Test Standards publish Psychological Testing: A Guide. This 17-page document offers guidance for non-psychologist test users in educational, clinical and occupational fields. It has three sections: an Introduction to testing, different applications and various quality control issues; Practical advice on what to look for in a test; Further information: where to go from here (which includes a short bibliography). The first section is a very brief textbook on testing, but the second section "amounts to a guide as to what to expect from a good test manual", and comes close (in four pages) to what ILTA is interested in identifying, I think.

==========

UK12: England: "The International Encyclopedia of Ed. Eval."

Record: UK12
TFTSmem: Alderson and Davidson
Country: England
Auth/Pub: Herbert J Walberg and Geneva D Haertel
Title(s): The International Encyclopaedia of Educational Evaluation. 1990. Oxford: Pergamon Press. 796 pp.
Language: English
Contact: The publisher: Pergamon Press plc, Headington Hill Hall, Oxford, OX3 0BW, England
Govt/Priv: priv
Lang/Ed: ed
StanDef: guideline

Comments:

Objective/Purpose:
[from p. xvii] "The International Encyclopedia of Education provides a current and comprehensive treatment of evaluation theories and practices, focusing especially on evaluation in

education. It is organized for effective use by both the beginning student of evaluation and the advanced practitioner both as a reference work and as a set of guiding articles for planning and conducting evaluation studies. The emphasis throughout is on practicality."

Target group/Audience intended:
There is no overt statement of the intended audience beyond that given immediately above. Judging from the content of the articles, the intended readership is probably quite wide. Articles are generally not overly technical.

Procedure:
This is an encyclopedia of articles on various topics in evaluation. Articles tend to average about five dual-column pages, with a upper length of ten pages.

Scope of influence:
Difficult to judge. It is a standard reference work which is likely found in many libraries and education agencies and companies.

Summary:
Following is a breakdown of the 164 articles which comprise this single-volume encyclopedia, given by the precise eight-'Part' headers as listed in the table of contents. There is also an 'Introduction' to each section authored by Walberg and Haertel.

Part 1. Evaluation approaches and strategies
- (a) Evaluation as a field of inquiry (5 articles)
- (b) Purposes and goals of evaluation studies (8 articles)
- (c) Evaluation models and approaches (9 articles)

Part 2. Conduct of and issues in evaluation studies
- (a) Normative dimensions of evaluation practice (4 articles)
- (b) Issues in test use and interpretation (10 articles)
- (c) Issues affecting sources of evaluation evidence (3 articles)

Part 3. Curriculum Evaluation
- (a) Models and philosophies (11 articles)
- (b) Components and applications of curriculum evaluation (9 articles)

Part 4. Measurement theory
- (a) General principles, models, and theories (10 articles)
- (b) Specalized measurement models and methods (12 articles)

Part 5. Measurement applications
- (a) Creation, scoring, and interpretation of tests (12 articles)
- (b) Using tests in evaluation contexts (11 Articles)

Part 6. Types of tests and examinations

(a) Test security, timing, and administration conditions (5 articles)
(b) Testing formats used in education (9 articles)
(c) Testing in educational settings (7 articles)
(d) Testing domains of knowledge, ability, and interest (12 articles) [contains one article on foreign language testing, authored by R.L. Jones]

Part 7. Research methodology
(a) Basic principles of design and analysis in evaluation research (7 articles)
(b) Issues in the design of quantitative evaluation studies (4 articles)
(c) Issues in qualitative evaluation research (3 articles)

Part 8. Educational policy and planning
(a) Evaluation research, decision making, social policy, and planning (7 articles)
(b) Dissemination and utilization of evaluation research (6 articles)

Commentary:
This seems an important work of scholarship. It is effectively an encyclopedia of educational assessment (generally) with great relevance to educational program evaluation (in particular). It should prove useful in all future ILTA standard-setting discussions.

==========

UK13: England: [BPS] "Code of Conduct Ethical Principles..."

Record: UK13
TFTSmem: Alderson
Country: England
Auth/Pub: The British Psychological Society
Title(s): Code of Conduct Ethical Principles and Guidelines, 1991
Language: English
Contact: Colin Newman, Executive Secretary, The British Psychological Society, St Andrews House, 48, Princess Road East, Leicester LE1 7DR, England. Fax: 01533 470787. e-mail: BPS1@le.ac.uk
Govt/Priv: priv
Lang/Ed: ed
StanDef: guideline

Comments:

Contains revised guidelines and codes of conduct.

==========

UK14: England: "Psychological Testing: A Practical Guide..."

Record: UK14
TFTSmem: Alderson
Country: England
Auth/Pub: J Toplis, V Dulewicz, and C Fletcher
Title(s): Psychological Testing: A practical Guide for Employers. Institute of Personnel Management 1987
Language: English
Contact: Colin Newman, Executive Secretary, The British Psychological Society, St Andrews House, 48, Princess Road East, Leicester LE1 7DR, England. Fax: 01533 470787. e-mail: BPS1@le.ac.uk
Govt/Priv: priv
Lang/Ed: ed
StanDef: guideline

Comments:

Contains an overview of testing: what tests are, what they can do, how to choose a test, how to use them, how to evaluate a testing programme.

==========

UK15: England: "Foreign Language Testing. Specialised..."

Record: UK15
TFTSmem: Alderson
Country: England
Auth/Pub: J L Trim and J A Price
Title(s): Foreign Language Testing. Specialised Bibliography 1, Second Edition, 1981. CILT
Language: English
Contact: Philippa Wright, Head of Information Services, Centre for Information on Language Teaching and Research (CILT) 20 Bedfordbury, London WC2N 4LB, England
Govt/Priv: priv
Lang/Ed: lang
StanDef: other

Comments:

The bibliography is divided into two parts: abstracts of 195 articles on testing over the years 1968-1981, and a selected list of 44 books and 20 published tests and syllabuses. The abstracts for articles are taken from the abstracting journal Language Teaching and Linguistics Abstracts. Nevertheless, the bibliography is less than comprehensive. It is of interest partly because it is one of the few such bibliographies and is often referred to by non-specialists, and partly because CILT considered it and the entry UK16 were of relevance. There is

very little directly of relevance to 'standards'

==========

UK16: England: "Foreign Language Testing Supplement..."

Record: UK16
TFTSmem: Alderson
Country: England
Auth/Pub: CILT
Title(s): Foreign Language Testing Supplement 1981-1987. Specialised Bibliography 6, 1988. CILT
Language: English
Contact: Philippa Wright, Head of Information Services, Centre for Information on Language Teaching and Research (CILT) 20 Bedfordbury, London WC2N 4LB, England
Govt/Priv: priv
Lang/Ed: lang
StanDef: other

Comments:

Is an update of UK15, of 200 articles, 4 survey articles, 11 bibliographies, 59 books, 6 journals/ newsletters, 38 articles not abstracted in Language Teaching, and 10 tests. Once again, the compilers used abstracts published in the same abstracting journal as UK15 and its successor, Language Teaching and the bibliography is far from complete. Contains a detailed subject and name index.

==========

UK17: England: "Testing Bibliography, Nov 1994. CILT"

Record: UK17
TFTSmem: Alderson
Country: England
Auth/Pub: CILT
Title(s): Testing Bibliography, Nov 1994. CILT
Language: English
Contact: Philippa Wright, Head of Information Services, Centre for Information on Language Teaching and Research (CILT) 20 Bedfordbury, London WC2N 4LB, England
Govt/Priv: priv
Lang/Ed: ed
StanDef: other

Comments:

A print out from the CILT catalogue of item on testing acquired since 1987 (76 items). No abstracts or annotations.

==========

UK18: England: "Info. Sheet 10: Guide to GCSE 'A' Level..."

Record: UK18
TFTSmem: Alderson
Country: England
Auth/Pub: CILT
Title(s): Information Sheet 10: Guide to GCE 'A' Level examinations; Information Sheet 37: GCSE examining groups. A Guide to language examinations on offer; Information Sheet 51: Guide to GCE 'AS' examinations; Information Sheet 56: Aternatives to GCSE and 'A' level examinations
Language: English
Contact: Philippa Wright, Head of Information Services, Centre for Information on Language Teaching and Research (CILT) 20 Bedfordbury, London WC2N 4LB, England
Govt/Priv: priv
Lang/Ed: ed
StanDef: test

Comments:

A useful set of lists of the main foreign language examinations in the UK, including graded tests. Contains brief descriptions of each examination or assessment, and how to get more information. Non-evaluative

==========

UK19: England and Wales: "GCE A and AS Code of Practice..."

Record: UK19
TFTSmem: Alderson
Country: England and Wales
Auth/Pub: Schools Curriculum and Assessment Authority
Title(s): GCE A and AS Code of Practice, July 1994
Language: English
Contact: C G Robinson, School Curriculum and Assessment Authority, Newcombe House, 45 Notting Hill Gate, London W11 3JE, England. Fax: 0171 221 2141
Govt/Priv: govt
Lang/Ed: ed
StanDef: guideline

Comments:

See UK05 for a related document, with similar structure. This is a key document for the UK. It is intended to "promote quality and consistency in the examining process across all

examining boards offering GCA Advanced (A) and Advanced Supplementary (AS) examinations. It will help to ensure that grading standards are constant in each subject and from year to year... The Code provides for the development of a system that seeks to ensure consistency, accuracy and fairness in the operation of all GCE A and AS examinations." It was jointly written by examining boards and has been adopted by them voluntarily. Boards are responsible for the implementation of the Code.

The Code is divided into eight sections, with two appendices on Principles for GCA A and AS examinations (dealing with syllabuses, assessment and reporting) and Special coursework arrangements, which stipulates the maximum % weighting for coursework in arrange of subjects, including Communication Studies, but not English or Modern Foreign Languages. The main sections are as follows:

1. Responsibilities of examining boards and examining board personnel
2. Syllabuses
3. Setting of question papers and provisional mark schemes for terminal examinations and end-of-module examinations
4. Standardisation of marking: terminal examinations and end-of-module examinations
5. Coursework assessment and moderation
6. Grading and awarding
7. The quality of language
8. Examining boards' relationship with centres.

It is impossible to give more detail without reproducing the document, but this is exactly the sort of document TFTS was hoping to find, and should be closely consulted in any discussion of ILTA standards.

==========

UK20: England and Wales: "Modern Foreign Languages..."

Record: UK20
TFTSmem: Alderson
Country: England and Wales
Auth/Pub: Schools Curriculum and Assessment Authority
Title(s): Modern Foreign Languages in the National Curriculum: Draft Proposals. May 1994
Language: English
Contact: C G Robinson, School Curriculum and Assessment Authority, Newcombe House, 45 Notting Hill Gate, London W11 3JE, England. Fax: 0171 221 2141
Govt/Priv: govt
Lang/Ed: lang
StanDef: performance

Comments:

This document is the result of a review of the National Curriculum by the new Chair of SCAA, Sir Ron Dearing. The review was intended to streamline the National Curriculum and contains detailed recommendations for this, in particular by "simplifying and clarifying the programmes of study, reducing the volume of material to be taught, reducing overall prescription so as to give more scope for professional judgment, and ensuring that the Orders are written in a way which offers maximum support to the classroom teacher." Essentially the document presents revised proposals for attainment targets, teaching learning activities and programmes of study, and is only indirectly relevant to the TFTS work, since it seeks to define standards of performance, not standards of practices.

==========

UK21: England: "Certificates in Commun. Skills in English ..."

Record: UK21
TFTSmem: Huhta
Country: England
Auth/Pub: University of Cambridge Local Examinations Syndicate
Title(s): Certificates in Communicative Skills in English (CCSE)

There are at least three documents about this test that seem to relate to standards:

a) CCSE: Administration and Centre Guidelines (1994) (19 pages)

b) CCSE: Examiners' Report.
- June 1991. (105 pages)
- November 1991. (101 pages)
- June 1992. (66 pages)
- November 1992 and June 1993. (216 pages)
- November 1993. (109 pages)

c) CCSE: Test of Writing. Samples of candidate performance on tasks from
- November 1990 and June 1991 papers (75 pages)
- June 1993 (61 pages)

In addition, the CCSE examination has a more general guidebook targeted to anyone interested in the content and format of the exam (CCSE: Examination Content & Administrative Information) which is not included in this summary.

Language: English
Contact: UCLES; EFL Division; Syndicate Buildings; 1 Hills Road; Cambridge CB1 2EU; UK. FAX: 01223-460278.
Govt/Priv: priv
Lang/Ed: lang
StanDef: test

Comments:

The CCSE is a general language examination for adults (16+) who intend to study or work in Britain. There are four levels (1-4) of tests in the four skills (reading, writing, listening, oral interaction). Testees are free to choose any combination of subtests at any of the four levels (e.g. only the test of reading at level 3 and the test of listening at level 2, but not the other subtests). The certificate reports the levels & tests passed.

Objectives and target groups:
a) To provide the test administrators (examination secretary, ushers, invigilators, interlocutors and assessors) with practical information on how to administer the tests.

b) To provide teachers preparing candidates for CCSE with practical examples of test performances and assessors' comments on them. Apparently test takers will find these reports useful, too.

c) As in b), but only on the writing tasks.

Procedure:
These documents are produced at the UCLES.

Scope of influence:
Apparently among the examination centres and the teachers who prepare candidates for the exam.

Summary:
a) Administration and Centre Guidelines

The document is a rather detailed guide and list of things that those involved in the practical administration of the exam have to do prior, during and after the administration. There are general considerations that the examinations secretary must take care of (entry forms, eligibility and withdrawals of candidates, handicapped candidates; what to do with the exam papers). This includes a checklist of the things to do at specific points of the administration process.

A lot of attention is devoted to running the oral interaction test which is a bit complicated procedure. Detailed advice is given on e.g. pairing the candidates, timetabling the pairs for different parts of the test (example of a timetable included), step by step guide for the usher for taking care of the candidates during the test (e.g. instructions, guiding

them to the right rooms), step by step guide for the interlocutor who interviews the candidates. The interlocutor's section also includes advice and examples on how to discuss with the candidates (what kind of questions are appropriate, how to extend the interaction; warnings against typical pitfalls). There is also some general information for the assessor.

Invigilators are also given advice on what to tell the candidates before and during the test session.

b) Examiners' Report:
These are published once or twice a year. Each document presents real test takers' responses for reading, writing, and listening tests (all tasks for the particular exam are presented). These examiners' comments on the responses are included (reasons for pass or fail). For the oral interaction test, there are obviously no authentic samples of performances; only the feedback collected from the examination centres is presented. The document starts with general information about the test, e.g. pass / fail rates for each level, as well as recommendations for candidate preparation for the teachers.

c) Samples of candidate performance on tasks from November 1990 and June 1991 & June 1993 papers.

These documents contain samples of test takers writings together with the assessors' comments. There is a brief general section on the criteria and marking of the writing tasks.

==========

UK22: England: "The IELTS Specifications..."

Record: UK22
TFTSmem: Wylie
Country: England
Auth/Pub: University of Cambridge Local Examinations Syndicate (UCLES), British Council, and IELTS Australia.
Title(s): The IELTS Specifications (International English Language Testing System). (draft) August 1994. (97 pages)
Language: English
Contact: Dr M. Milanovic, UCLES, 1 Hills Road, Cambridge CB1 2EU, UK. FAX: 01223-460278.
Govt/Priv: priv
Lang/Ed: lang
StanDef: guideline

Comments:

This document is listed as a UK record although it has influence in Australia and New Zealand and internationally as well.

Objective:
To inform about the history, purpose, administration, theoretical models underpinning, general content, and code of practice for the IELTS test, as well as stating specifications for the various modules.

Intended audience/target group(s):
IELTS Advisory Committee, chief examiners, item-writers, question moderators and researchers. (IELTS is a secure test, used largely for determining readiness of overseas students of non-English-speaking background for tertiary studies in Australia, NZ and UK, and the document has a restricted circulation).

Procedure:
Developed by officers at UCLES, with input from the IELTS Advisory Committee.

Scope of influence:
To be observed by all those involved in developing and trialling IELTS tests, assessing performance and reporting results. It also states that all groups with which UCLES works are encouraged to "...adopt the policies on confidentiality stated in the Code of Practice (Chapter 11), so that data transferred from them or to them is adequately protected."

Summary:
As well as chapters on test purpose and administration, models of language ability, test content, aspects of linguistic and strategic competence, and the IELTS Code of Practice, there is a separate chapter for the specifications of various sub-tests - Listening, Academic Reading, General Training Reading, Academic Writing, General Training Writing, and Speaking.

Commentary:
The Code of Practice outlines systems and procedures for validating the IELTS test, evaluating the impact of the test, providing relevant information to test users, and ensuring that a high quality of service is maintained. The final section includes the following commitment as part of quality of service: "Supporting the activities of professional associations involved in developing and implementing professional standards and codes, making available the results of research, and seeking peer review of its activities".

==========

UK23: Scotland: [Brochures:] "National Certificate and You..."

Record: UK23
TFTSmem: Alderson
Country: Scotland
Auth/Pub: Scottish Vocational Education Council

Title(s): Brochures: National Certificate and You, Stepping Stones to Success, and A Guide for Teachers
Language: English
Contact: Chris Brown, Assistant Director, Research, SCOTVEC, Hanover House, 24 Douglas Street, Glasgow, G2 7NQ
Govt/Priv: govt
Lang/Ed: ed
StanDef: guideline

Comments:

The latter (A Guide...) includes a brief description of a quality assurance scheme used to monitor qualifications at school and national levels. For example, it says that schools "have to ensure that standards within the school are consistent and must involve regular meetings of staff who are teaching and assessing the same modules to: agree the design of assessment so that all pupils taking the same module receive comparable assessment; agree their marking standards so that there is internal consistency across all pupils; review the results and agree on certification decisions. "External verifiers visit schools "to make sure that national standards are being maintained". Such verifiers may ask to see: "records of achievement for pupils; assessment instruments used to cover a module's outcomes; annotated assessment evidence for each outcome and group of pupils". No further details or standards are available.

In addition, numerous documents describing the detailed specifications of the National Certificate Modules in a range of different languages. Relates to the Standards of the Languages Lead Body (UK08). Reference is made to documents not yet seen: SCOTVEC's National Standards for Assessment and Verification, Guidelines for Module Writers, and SCOTVEC's Guide to Assessment

==========

UK24: Scotland: [Scot. Exam. Board: Various documents]

Record: UK24
TFTSmem: Alderson
Country: Scotland
Auth/Pub: Scottish Examinations Board
Title(s): Various documents relating to specifications and sample examination papers
Language: English
Contact: V M Kelly, External Relations, Scottish Examinations Board, Ironmills Road, Dalkeith, Midlothian, EH22 1LE Fax: 0131 654 2664
Govt/Priv: govt
Lang/Ed: ed
StanDef: performance

Comments:

Details of Subject Arrangements: Specifications for the assessment of foreign languages, including some mention of assessment criteria and timing of examinations. Also enclosed Sample question papers for examinations in numerous languages at Higher Grade and Certificate of Sixth year Studies

==========

UK25: Scotland: "Communicative Language Testing: a Resource..."

Record: UK25
TFTSmem: Alderson (does not have a copy)
Country: Scotland
Auth/Pub: P S Green
Title(s): Communicative language testing: a resource handbook for teacher trainers. Council for Cultural Cooperation, Project No 12. Council of Europe, 1987 (ISBN 92 871 1052 2)
Language: English
Contact: Graham Thorpe, Research Services Unit, The Scottish Council for Research in Education, 15 St John Street, Edinburgh, EH8 8JR, Scotland
Govt/Priv: priv
Lang/Ed: lang
StanDef: other

Comments:

"Contains some practical examples but is largely theoretical covering discussion at the workshop". ILTA does not hold a copy

==========

UK26: Scotland: "Curriculum and Assessment in Scotland..."

Record: UK26
TFTSmem: Alderson
Country: Scotland
Auth/Pub: The Scottish Office Education Department
Title(s): Curriculum and Assessment in Scotland: National Guidelines. i) Assessment 5-14 (October 1991) ii) English Language 5-14, June 1991 iii) Reporting 5-14: Promoting Partnership, Nov 1992 iv) The Structure and Balance of the Curriculum 5-14, June 1993 v) Gaelic 5-14, June 1993
Language: English
Contact: Graham Thorpe, Research Services Unit, The Scottish Council for Research in Education, 15 St John Street, Edinburgh, EH8 8JR, Scotland
Govt/Priv: priv

Lang/Ed: ed
StanDef: guideline

Comments:

These documents are issued by the Scottish Office, and represent Government policy with respect to assessment in Scotland. They are intended to convey to education authorities guidance on the principles which "should underlie school policies for the assessment of pupils' progress and attainment.....The guidance provides a sound basis for effective, coherent and manageable assessment of pupils' achievements in relation to standards of attainment set out in the 5-14 curriculum guidelines. It also recognises the role of national test as a means of confirming classroom based assessment and of conveying information on progress to parents. Using these guidelines, schools should be able to develop effective assessment policies and practices which will improve the quality of learning and teaching."

Document i) introduces the main features of a strategy for assessment in the context of teaching and learning, and offers guidance on the basis of which schools should review and develop their assessment policy. It describes the principles and intentions which should underlie each schools' assessment policy in the areas of planning, teaching, recording, reporting, evaluating. It also discusses who will be involved in assessment. Parts 1 and 2 are available to all primary and secondary teachers. Part 3 to this document is published separately and made available to all primary and secondary schools (see document UK27 below). It is a very useful document on good assessment practice for classroom teachers, rather than a set of standards on the lines of the APA, and therefore of considerable interest to ILTA

Documents ii) and v) give details of the attainment outcomes and targets, and the programmes of study for the two languages, representing in effect a national curriculum for Scotland which parallels the English curriculum in structure if not in detailed content. There is additional information on assessment and recording, and specific issues in areas like Knowledge about Language, Culture, Genre, Mass Media and the like, which are not directly related to assessment.
Document iii) provides guidance to teachers and schools on what makes a good report, report formats, frequency of reporting to parents, how to complete the report, involve others and how to use the report. Special issues to do with equality of opportunity (gender, special needs, multicultural and social issues) are also addressed. The document presents recommendations and details of good practice, for teachers and schools and, again, is of relevance to classroom based assessment for ILTA.

Document iv) provides guidance on the structure and balance of

Scotland's curriculum in the light of developments since 1989. A brief Appendix gives the main details of the Attainment Outcomes in the main curricular areas, and is of only tangential interest to ILTA.

==========

UK27: Scotland: "Assessment 5-14. Improving the Quality..."

Record: UK27
TFTSmem: Alderson
Country: Scotland
Auth/Pub: Committee on Assessment, Scottish Education Department
Title(s): Assessment 5-14. Improving the Quality of Teaching and Learning. Part 3: A Staff Development Pack. The Scottish Education Department, September 1990
Language: English
Contact: Graham Thorpe, Research Services Unit, The Scottish Council for Research in Education, 15 St John Street, Edinburgh, EH8 8JR, Scotland
Govt/Priv: priv
Lang/Ed: ed
StanDef: guideline

Comments:

This document is intended to accompany the documents in UK26, and provides ideas, advice and staff development materials to help schools and teachers find "manageable ways of assessing effectively, and to set assessment in the wider context of learning and teaching." The document is organised under five headings: Planning, teaching, Recording, Reporting and Evaluating. It includes discussions of teacher judgments and on what to base them, on obtaining evidence, and on recording and acting on that evidence. Its remit is of course much broader than testing, although elements are of relevance.

==========

UK28: Scotland: "Taking a Closer Look. A Resource Pack..."

Record: UK28
TFTSmem: Alderson
Country: Scotland
Auth/Pub: Schools' Assessment Research and Support Unit, (SCRE)
Title(s): Taking a Closer Look. A resource pack for teachers. Assessment 5-14.
Language: English
Contact: Graham Thorpe, Research Services Unit, The Scottish Council for Research in Education, 15 St John Street, Edinburgh, EH8 8JR, Scotland

Govt/Priv: priv
Lang/Ed: ed
StanDef: guideline

Comments:

A fairly brief document giving teachers advice on diagnostic procedures in assessment, setting out a number of clear principles for effective learning, teaching and assessment, and introducing a number of publications giving details of diagnostic procedures for mathematics, English language and science. The English language series "Taking a Closer Look at English Language" is said to focus on diagnostic procedures in Writing. A volume on Reading is said to be under preparation.

==========

UK29: Scotland: [Scot. Exam. Board, Cmte. on Testing: various]

Record: UK29
TFTSmem: Alderson
Country: Scotland
Auth/Pub: Committee on Testing, Scottish Examination Board
Title(s): i) Assessment 5-14: A Teacher's Guide to National Testing in Primary Schools, 1993 ii) Assessment 5-14: A Teacher's Guide to National Testing in Secondary Schools, 1993 iii) The Framework for National Testing iv) Catalogue of National Test Units, 1994 v) National Tests in Reading: Information for Teachers, 1994 vi) National Tests in Writing: Information for Teachers, 1994
Language: English
Contact: The 5-14 Assessment Unit (FFAU), The Scottish Examination Board, Ironmills Road, Dalkeith, Midlothian EH22 1LE
Govt/Priv: govt
Lang/Ed: ed
StanDef: guideline

Comments:

These documents accompany those in UK28. Documents i) and ii) give an introduction to National testing in Scotland, which applied in all primary schools since 1993, and in secondary schools since 1994. They set out the purpose of National Tests, provide guidance on their use within teaching and learning, and give details on who should do what when. Document iii) is divided into sections: the purpose of the tests, key principles of national testing, characteristics of the National tests, how they are structured, how schools choose the tests, when pupils take them, which pupils will be tested, how they will be ordered, administered, marked, records kept, marks converted to levels, how the testing will

be monitored, and how the results will be used. Although brief, this document is an example of 'good practice' in test use and therefore relevant to ILTA.

Document iv) describes in considerable detail the testing units available for testing Language and Mathematics for Levels A-E, within the National testing framework. A very useful appendix contains answers to questions most frequently asked concerning National Testing, including "Will my school be moderated?", "Why are some schools asked to pre-test units?" "Are the tests confidential?" (The answer is No!) Document v) and vi) give teachers detailed and practical guidance on the administration and marking of national tests and units of reading and writing, threshold scores of units and unit record sheets. In addition, document vi) gives detailed criteria for marking and how to use them

==========

US01: U.S.A.: [APA/AERA/NCME] "Standards for Ed. & Psych. ..."

Record: US01
TFTSmem: Davidson and Douglas (The 1985 version is available at many libraries and is widely used in college-level training in educational measurement. Davidson and Douglas have copies of all the other documents dated below.)
Country: U.S.A.
Auth/Pub: American Psychological Association (APA) / American Educational Research Association (AERA) / National Council on Measurement in Education (NCME)
Title(s): Standards for Educational and Psychological Testing (1985, 1974, 1966; precursor documents: 1955, 1954; historical antecedents dating back to 1890). The bulk of the remarks under 'Comments' below refer to the current 1985 edition, in acknowledgement of AERA/APA/NCME's ongoing wish for each publication to supersede the previous. This desire to supersede previous versions is probably why there are no "edition" or "volume" numbers on the 1974 and 1985 documents.
Language: English
Contact: American Psychological Association, 1200 Seventeenth St., NW, Washington, DC 20036 USA, or many libraries.
Govt/Priv: priv
Lang/Ed: ed
StanDef: guideline

Comments:

Objective/Purpose:
The objective and purpose of this document is well-summarized in

its Introduction:

> "Although not all tests are well-developed, nor are al testing practices wise and beneficial, available evidence supports the judgment of the Committee on Ability Testing of the National Research Council that the proper use of well-constructed and validated tests provides a better basis for making some important decisions about individuals and programs than would otherwise be available.
>
> Educational and psychological testing has also been the target of extensive scrutiny, criticism, and debate both outside and within the professional testing community. The most frequent criticisms are that tests play too great a role in the lives of students and employees and that tests are biased and exclusionary. In consideration of these and other criticisms, the Standards is intended to provide a basis for evaluating the quality of testing practices as they affect the various parties involved." (p.1)

The Preface provides further background on the intent of these the joint committee of these three agencies which authored the Standards:

> "The Standards should:
>
> 1. Address issues of test use in a variety of applications.
>
> 2. Be a statement of technical standards for sound professional practice and not a social action prescription.
>
> 3. Make it possible to determine the technical adequacy of a test, the appropriateness and propriety of specific applications, and the reasonableness based on the test results.
>
> 4. Require that test developers, publishers, and users collect and make available sufficient information to enable a qualified reviewer to determine whether applicable standards are met.
>
> 5. Embody a strong ethical imperative, though it was understood that the Standards itself would not contain enforcement mechanisms.
>
> 6. Recognize that all standards will not be uniformly applicable across a wide range of instruments and uses.
>
> 7. Be presented at a level that would enable a wide range of people who work with tests or test results to use the Standards.
>
> 8. Not inhibit experimentation in the development, use,

and interpretation of tests.

9. Reflect the current level of consensus of recognized experts.

10. Supersede the 1974 Standards for Educational and Psychological Tests." (p. v)

Target group/Audience intended:
Based on the Introduction, the authors of this document intend it to be used by the following individuals: test developer, test user, test taker, test sponsor, test administrator, and test reviewer. (p.1)

Procedure:
Produced by a joint committee of AERA, APA and NCME. Currently under revision to produce a fourth edition.

Scope of influence:
Quite wide. See 'Commentary' below.

Summary:
The document is a presentation of a number of standards of one to several sentences in length, arranged in chapters and parts. Each chapter includes a discussion of 'background' related to its topic. Standards are categorized as being of primary, secondary, or conditional "importance [which is] viewed largely as a function of the potential impact that the testing process has on individuals, institutions, and society." (p. 2). Following is the Table of Contents, annotated with the number of standards present in each chapter:

Part I: "Technical Standards for Test Construction and Evaluation"

Chapter 1: "Validity" (25 Standards)
Chapter 2: "Reliability and Errors of Measurement"(12 Standards)
Chapter 3: "Test Development and Revision" (25 Standards)
Chapter 4: "Scaling, Norming, Score Comparability, and Equating" (9 Standards)
Chapter 5: "Test Publication: Technical Manuals and User's Guides." (11 Standards)

Part II: "Professional Standards for Test Use"

Chapter 6: "General Principles of Test Use" (13 Standards)
Chapter 7: "Clinical Testing" (6 Standards)
Chapter 8: "Educational Testing and Psychological Testing in the Schools" (12 Standards)
Chapter 9: "Test Use in Counseling" (9 Standards)
Chapter 10: "Employment Testing" (9 Standards)
Chapter 11: "Professional and Occupational Licensure and

Certification" (5 Standards)
Chapter 12: "Program Evaluation" (8 Standards)

Part III: "Standards for Particular Applications"

Chapter 13: "Testing Linguistic Minorities" (7 Standards)
Chapter 14: "Testing People Who Have Handicapping Conditions" (8 Standards)

Part IV: "Standards for Administrative Procedures"

Chapter 15: "Test Administration, Scoring, and Reporting" (11 Standards)
Chapter 16: "Protecting the Rights of Test Takers" (10 Standards)

Glossary
Bibliography
Index

Commentary:

Introduction

A powerful force, if not the key 'player', in the U.S. assessment guideline scene is the document published by the American Educational Research Association, American Psychological Association and the National Council on Measurement in Education (1985), entitled Standards for Educational and Psychological Testing. We contend that the letter of those Standards are influential in most educational assessment practice in the U.S.A., language testing being no exception. If the letter of those guidelines is not followed, then their spirit often is, and if a test follows neither the spirit nor the letter of the Standards, then criticism often reflects principles of the Standards.

In these comments, we first review the historical evolution of the Standards and a related development in U.S. testing: the test bibliography industry (e.g. Buros' Mental Measurement Yearbook or Tests In Print series, among others). To illustrate the force which the document has in U.S. society, we then present examples of the relevance of the Standards to educational measurement in the USA, drawn from court litigation, statutes and regulations. In closing, we attempt to evolve a definition of 'standards' that best represents the U.S. educational scene.

Historical Evolution of the Standards

It is clear that the Standards originated in historical work on 'standard measures', or the desire by early U.S. psychol-

ogists to have common agreed-upon measures for human traits. This desire began toward the end of the last century, and was a productive debate for about two decades. That debate subsided, and it was not until the 1950s that the predecessors of the present Standards were written, in which the notion 'standard' bore a different meaning: a guideline for good practice. We contend that this re-definition was natural and expected, given the evolution and growth of the test-evaluation industry, led primarily by Oscar Buros' Mental Measurement Yearbook series.

The genesis of the APA/AERA/NCME document reaches back 100 years or more, to the end of the previous century. James McKeen Cattell argued in 1890:

> Psychology cannot attain the certainty and exactness of the physical sciences, unless it rests on a foundation of experiment and measurement. A step in this direction could be made by applying a series of mental tests and measurements to a large number of individuals. The results would be of considerable scientific value in discovering the consistency of mental processes, their interdependence, and their variation under different circumstances. Individuals, besides, would find their tests interesting, and perhaps, useful in regard to training, mode of life, or indication of disease. The scientific and practical value of such tests would be much increased should a uniform system be adopted, so that determinations made a different times and places could be compared and combined. (Cattell, 1890: 373)

In that article, Cattell provided his recommendations for common measures of ten types of human behavior.

In the early 1890s, the American Psychological Association was formed. In its fourth year, the Association appointed a "Committee on Physical and Mental Tests" which Cattell chaired. That committee reported to the APA at its Fifth Annual Meeting in 1896 (APA, 1897). The report was similar in structure to Cattell's 1890 paper, in that it listed a number of recommended measures for certain human traits. Near the end of that report was the following comment, which seems to foreshadow the organization of the present Standards: "[The committee] does not recommend that the same tests be made everywhere, but, on the contrary, advises that, at the present time, a variety of tests be tried, so that the best ones may be determined." (APA, 1897: 137). Thereafter, the Committee faded in APA records, though as noted below, in the 1954 precursor to the present-day Standards there is reference a 1906 committee. This early interest in testing was short-lived, but clearly, it defined its mission as the location of 'standards' where that term denotes standard or accepted measures. Cattell's 1890 paper included, for example, the following paragraph, which was based on his work at a lab at the University of Pennsylvania:

> Memory and attention may be tested by determining how many letters can be repeated on hearing once. I name distinctly and at the rate of two per second six letters, and if the

> experimentee can repeat these after me I go on to seven, then eight, etc.; if the six are not correctly repeated after three trials (with different letters), I give five, four, etc. The maximum number of letters which can be grasped and remembered is thus determined. Consonants only should be used in order to avoid syllables. (Cattell, 1890: 377)

The above procedure is strikingly similar to modern adaptive testing, though admittedly, it is more focused on a rather narrow skill.

In the middle and late 1930s, Oscar Buros of Rutgers University formed an initiative which, on the surface, appears to continue the definition of 'standard' as 'standard test' begun by Cattell and the APA Committee. Buros' first efforts in this regard were a series of articles published in the 1930s. These were fairly short overviews of some common psychological tests, with bibliographic information provided. Buros, with support from Rutgers, then expanded the project to provide a critical summary of each test in his bibliography, and the result was the first edition of The Mental Measurement Yearbook (MMY). For more information on the MMY, see record US03. Some years later, Buros launched a second publication, Tests In Print (TIP). This book is a much more concise bibliography of information about tests. It does not include critical summary or comment about each test, but it does cross-reference the user to editions of the MMY where such commentary can be found. See record US04 for more information on TIP.

What is notable about this test bibliography industry is that it essentially took up the trend started by Cattell in the 1890s, that of presenting a reference work to a number of tests. In the 1890s, that work was primarily a list of procedures, such as Cattell's recipe for the measurement of memory, above. By the start of Buros' work in the 1930s, there were enough professionally developed and commercial tests that Buros could provide entries organized by the tests themselves.

It is fair to say that by the early 1950s, educational and psychological testing was a large concern in the USA. Particularly in the post-World War II boom in education, the demand for access to learning made it necessary to assess large numbers of people quickly and efficiently. Texts on educational measurement were appearing in the 1930s, with some large and influential books coming to dominate, e.g. Lindquist, 1951. By the early 1950s, testing was an industry producing many important products that were indexed in the MMY and which were guided by practical tools such as Lindquist's text. It is not surprising, then, that in 1954 the APA published a supplement to The Psychological Bulletin. This document is the direct genesis of the present-day Standards, as noted in the 'Title(s)' field, above.

In 1954, the APA wrote the following, which confirms the above reasoning about the efforts of some fifty years earlier and sets forth the mandates which have guided the present-day Standards (see 'Objectives/Purpose' above):

> In 1906, an APA committee, with Angell as chairman, was

> appointed to act as a general control committee on the subject of measurements. The purpose of their work was to standardize testing techniques, whereas the present effort is concerned with standards of reporting information about tests. (APA, 1954: 1)

While we have been unable to uncover direct evidence, it is possible that the APA was influenced by the MMY and felt that the MMY fulfilled much the same mission as that of the APA committees of the period 1890-1910. In any regard, by 1954, the APA saw as its charge 'standards' of practice in measurement, and not the endorsement of standard measures. We contend that this was a critical juncture in the history of standard-setting in assessment in the USA. By the mid-1950s, two streams had emerged:

> (1) 'Standard tests', to the extent that such a concept is reasonable, were indexed in the MMY and later documents in the test bibliography industry.
>
> (2) Standards of practice were put forth by the APA/AERA/NCME.

In 1955, the American Educational Research Association (AERA) and the National Council on Measurement in Eduation (NCME) published their own document, which was quite similar to the 1954 APA publication (AERA/NCME, 1955). The Foreword of the 1955 document cites the 1954 APA publication:

> The [AERA/NCME] Committee profited materially from the work done previously by the Committee on Test Standards of the American Psychological Association ... which appeared ... [in] 1954". (AERA/NCME, 1955: 3)

The AERA/NCME document closely followed the layout of the 1954 APA publication, so it is not surprising that the three agencies soon formed a joint committee. In 1966, that joint APA/AERA/NCME committee issued Standards for Educational and Psychological Tests and Manuals, which is the first edition of the present Standards. A second edition was published in 1974. The present third edition was published in 1985, and the APA, AERA and NCME are at work on a fourth edition as of this writing. Given that a fourth APA/AERA/NCME document is in the works, a trend of revision every decade is being established. The order of authorship changed with the 1985 document, as shown in the reference list for US01, below.

All these publications from 1954 to 1985 reflect a definition of 'standards' to mean guidelines of good practice. In particular, each successive edition of the triple-authored Standards (1966, 1972, 1985) provide an interesting historical record of the evolution of the scope and depth of such standards. Content, organization, and formatting changed over those 19 years; for example, the 1985 Standards added a section on the testing of linguistic minorities and embraced a unified definition of test validity.

The Standards and The Law

Recently, the APA/AERA/NCME Standards have become cited in legal decisions, U.S. federal non-statutory regulations, and in one state are cited directly in statute.

In the case of Watson vs. Fort Worth Bank and Trust, argued before the U.S. Supreme Court in 1988, the Standards were instrumental in the Court's decision to "vacate and remand". Such a decision means that the Court heard the case but decided that the immediately lower court (from which the case was referred) did not adequately try the case. Ms. Watson, the plaintiff, had alleged racial discrimination in a promotion dispute at the Bank which involved certain subjective promotion assessments. The APA prepared an *amicus curiae* (Friend of the Court) brief in support of Watson, and the Court cited that brief and the Standards in its decision as support for remanding to the lower court. The APA brief argued, among other points, that subjective measures should be held to rigorous technical development to the same extent as objective measures, and it is that point which the Court cited. Perhaps APA hoped for a Supreme Court decision in favor of Ms. Watson, which in the culture of U.S. Supreme Court decisions is a feather in the *amicus curiae* organization's cap. Instead, the Court did not decide either for Watson or the Bank, but rather to send the matter back to the immediately lower court, saying that the very point which the APA cited was not adequately tried at that lower level. This resulted in a partial victory for the Standards' role in the U.S. legal system, at best. (See 487 U.S. 977 and amicus curiae brief filed by the APA, U.S. Supreme Court docket number 86-6139 [microfiche]).

The Standards have presence elsewhere in the U.S. legal system. In the case of Richardson vs. Lamar City Board of Education (of Alabama), a case involving alleged racial discrimination in tenure and promotion in a secondary school, the judge decided in favor of the plaintiff. He cited the Standards extensively in his written judgment. (See 729 F. Supp. 806). We also see the Standards cited in non-legislative (and less binding) regulation at the federal level; they are cited by the Federal Equal Opportunity Employment Commission, or EEOC. The EEOC is an oversight agency which guards against discrimination by employers who receive federal money, as for example a university might receive. (See Code of Federal Regulations Vol. 29, Part 1607.5 [1985]). Finally, the Standards are cited in actual legislative statute in one state. They appear in the education law of the State of South Carolina, in a section dealing with the mandate to teach and assess higher order skills in elementary and secondary schools:

> When selecting nationally normed achievement tests for the statewide testing program, the State Board of Education shall endeavor to select tests with a sufficient number of items

> which may be utilized to evaluate students' higher order thinking skills. The items may be used for this purpose only if the test created from the items meets applicable criteria set forth in the American Psychological Association publication "Standards for Educational and Psychological Testing" (Code of Laws of South Carolina 1976. Vol. 20: Education. Rochester NY: The Lawyer's Co-Operative Publishing Company. Enacted 1989, Act No. 194, sec. 13)

In the USA, there are two types of law. The first is statutory law and non-statutory regulation. Statutory law has the weight of legislative enactment behind it. The South Carolina law above is an example. State and Federal regulations, usually written by a government agency whose task it is to write such documents, do not have legislative force, but they do exert some governance in civic matters and are part of the concept of statutory law. The second type of U.S. law is common law, or the evolving system of judicial decisions that come from actual cases tried in court. The Richardson case could become part of common law. The Watson decision is an even better example, because the U.S. Supreme Court is the ultimate and final appelate body in the country. Either the Richardson or Watson decisions can be cited in later legal action, and such citation can lend common law power and credence to the authority of the Standards. Granted, the fact that the Supreme Court remanded the Watson decision makes creates less of a precedent than if the Court had decided for Watson. Had that happened, then the Standards might have achieved a legal status similar to that of other, more famous Supreme Court decisions (such as those on abortion or desegregation) which have come to have the force the highest common law possible in the USA: a Supreme Court decision.

Implicit in the U.S. legal system is the notion of precedent. Common law is essentially a matter of precedent, and common law decisions can become incorporated in later legislative statute, as for example the desegregation decisions in the 1950s which influenced civil rights legislation about a decade later.

We now offer a definition of human measurement 'standards' from a U.S. perspective, a definition which incorporates both the legal state of affairs and the historical precedents which helped to generate it.

'Standards' in U.S. Educational and Psychological Measurement

We now offer a definition of 'standards' based on our analysis in this record:

> Standards of educational and psychological measurement are authoritative guidelines for good practice or particular widely accepted practices or measures.

This definition embraces both the meaning of 'standard' as 'standard measure' and the meaning which connotes technical and

professional quality assurance -- the former is the trend started by the APA committees at the turn of the century and continued by the MMY and similar publications, and the latter is the trend established by the APA in 1954 and continuing today in the fourth revision of the APA/AERA/NCME Standards. Although we have some interesting legal records cited above, we do not claim that the Standards have, in fact, achieved the status of common or statutory law in the USA. They are entering that domain as an influence.

Using as a database all the above cited materials, we do not see wide acceptance in the USA of a third meaning of 'standard' to mean 'cut score' (Alderson et al., 1995: Chapter 11). In this third meaning, a standard is a certain score or proficiency level at which some decision is made or certification awarded.

Finally, we should note that it is possible that the definition of 'standard', to mean 'standard measure' may be frowned upon by the spirit if not the letter of the current Standards. On reading those documents from 1954 to 1985, it is clear that the APA/AERA/NCME are far more concerned that tests are developed properly than they are worried about endorsing any particular measure. Perhaps this reflects the vastly decentralized nature of psychology and education in the USA; we are a nation without a strong governmental 'Ministry of Education', 'Department of Human Research' or the like to govern psychological research. Elemental civics dictate human measurement in the USA: capitalistic choice from a large array of alternatives. It is unlikely that a U.S. court case could turn upon failure to use a certain widely-used test, such as the TOEFL, though it is quite interesting that we did locate some court cases above which cited the APA/AERA/NCME Standards.

That said, we still note that the notion of a standard measure is not anathema to the U.S. educational scene; for that reason, we include it in our definition above. We need only cite the influence of certain powerful tests here, such the large use of the TOEFL for entry decisions about international college students who come here to study. Perhaps such tests are not called 'standard' by their users; perhaps they are. But they have, we contend, achieved a certain status based on widespread acceptance, not unlike the measures which Cattell and his colleagues tried to define some one hundred years ago.

References cited in Record US01:

(Note: legal references, in keeping with legal tradition, are given in the body of the commentary above and not reprinted here.)

(Historical note: the evolution of the present day 1985 Standards is as follows: APA/AERA/NCMUE (1954) --> AERA/NCMUE (1955) --> APA/AERA/NCME (1966) --> APA/AERA/NCME (1974) --> AERA/APA/NCME (1985). Full references are given below in author order.)

Cattell, J.M. 1890. Mental tests and measurements. Mind (15), pp. 373-381.

American Psychological Association. 1897. Preliminary Report of the Committee on Physical and Mental Tests [reported in the] Proceedings of the Fifth Annual Meeting of the American Psychological Association, Boston [and Cambridge], December 1896. The Psychological Review 4:2, pp. 132-138.

American Psychological Association / American Educational Research Association / National Council on Measurements Used in Education [former name of NCME]. 1954. Technical recommendations for psychological tests and diagnostic techniques. Supplement to The Psychological Bulletin. 52:2, Part 2. pp. 1-38.

American Psychological Association / American Educational Research Association / National Council on Measurement in Education. 1966. Standards for Educational and Psychological Tests and Manuals. Washington, DC: American Psychological Association.

----------. 1974. Standards for Educational and Psychological Tests. Washington, DC: American Psychological Association.

American Educational Research Association / National Council on Measurements Used in Education [former name of NCME]. 1955. Technical Recommendations for Achievement Tests. Washington, DC: National Education Association.

American Educational Research Association / American Psychological Association / National Council on Measurement in Education. 1985. Standards for Educational and Psychological Testing. Washington, DC: American Psychological Association.

Lindquist, E.F. (Ed.) 1951. Educational Measurement. Washington, DC: American Council on Education

==========

US02: U.S.A.: "ETS Standards for Quality and Fairness"

Record: US02
TFTSmem: Douglas
Country: U.S.A.
Auth/Pub: Educational Testing Service
Title(s): ETS Standards for Quality and Fairness
Language: English
Contact: Carol Taylor, TOEFL 2000 Project, ETS, Princeton NJ, USA

Govt/Priv: priv
Lang/Ed: ed
StanDef: guideline

Comments:

Reviewed in Alderson, Clapham and Wall. (forthcoming) Language Test Construction and Evaluation. Cambridge UP.

Objectives/Purpose:
To establish standards for good testing practice within ETS; might not be applicable outside that context.

Target group/Audience intended:
Testing professionals, not the general public: ETS item writers, test assemblers, analysts, user services personnel.

Procedures:
The Standards represent corporate policy and were produced by a wide spectrum of ETS staff and administrators, in collaboration with consultants from outside the corporation. The Standards are based on the AERA/APA/NCME standards, as interpreted within the ETS context. They include a regulatory mechanism in the form of an ETS Office of Corporate Quality Assurance, and a "Visiting Committee" made up of outside consultants to monitor compliance with the Standards.

Scope of Influence:
Limited to the ETS organization, although the Standards are available to testers outside the corporation.

Summary:
The Standards cover seven areas: Accountability, Confidentiality of Data, Quality Control for Accuracy and Timeliness, Research and Development, Tests and Measurement, Test Use, and Public Information. The guidelines under each section are detailed and comprehensive. The Standards document contains a comprehensive Glossary to clarify key terms.

Commentary:
The Standards establish a model for good professional/commercial language testing practice. The fact that they are enforceable within the organization (e.g. programs and funding within ETS can be cut off for failure to abide by the Standards) is perhaps, in addition to their thoroughness, their most salient feature. The Standards are certainly worth studying by those who wish to establish a set of guidelines for good testing practice. Certainly the sections on Research and Development, Tests and Measurement, and Test Use are generalizable to non-ETS contexts.

==========

US03: U.S.A.: "Mental Measurement Yearbook"

Record:	US03
TFTSmem:	Douglas (This publication is also available at many libraries)
Country:	U.S.A.
Auth/Pub:	Buros Institute
Title(s):	Mental Measurement Yearbook
Language:	English
Contact:	Buros Institute of Mental Measurements, University of Nebraska, Lincoln NE 68588, USA
Govt/Priv:	priv
Lang/Ed:	ed
StanDef:	test

Comments:

(see also record US01)

Objective/Purpose:
"To impel test authors to publish higher quality tests with detailed information on their validity and limitations; to foster in test users a greater awareness of both the values and limitations involved in the use of standardized tests; to stimulate test reviewers and others to consider more thoroughly their own values and beliefs in regard to testing; to suggest more discerning methods to test users of arriving at their own appraisals of tests in light of their particular values and needs; and to make test users aware of the importance of being suspicious of all tests - even those produced by well-known authors and publishers - which are not accompanied by detailed data on their construction, validation, uses, and limitations" (1978 MMY).

Target group/Audience intended:
"Designed to assist test users in education, psychology, and industry to choose more discriminatingly from the many tests available."

Procedure:
Only tests which are new or revised since the last MMY are included. Reviews are written by "well qualified professional people who were selected by the editors on the basis of their expertise in measurement and, often, the content of the test being reviewed" (1992 MMY).

Scope of influence:
A standard, well-known reference work.

Summary:
1992 MMY contains 703 reviews of 477 commercially available tests for use by English speaking subjects, including measures of personality, vocational interest, reading intelligence, mathematics, speech & hearing, English, science, social studies, achieve-

ment, fine arts, sensory motor, and multi-aptitude. Reading tests account for 4.4% of the entries, speech & hearing for another 4.2%. There are only 8 entries dealing with English as a foreign language, and none on other foreign language tests.

Each entry contains information about ordering and costs, a summary of the intended uses for the test, manuals and accompanying materials, and one or more critical reviews.

Commentary:
First published in 1938, the most recent edition is the Eleventh, published in 1992. The next edition is due out in 1995. The recent editions are much smaller in scope than earlier ones: the 1978 edition was in two volumes and contained reviews of some 1800 tests. As a result of the more manageable size nowadays, the entries are more up to date. However, owing to the still vast size of the undertaking, the quality of the reviews varies quite a bit. Still it is a good first reference for test users just getting started in choosing an instrument. Earlier volumes can be a good source of information for studies of the development of mental measurement and changing views of good testing practice.

==========

US04: U.S.A.: "Tests in Print"

Record: US04
TFTSmem: Douglas (This publication is also available at many libraries)
Country: U.S.A.
Auth/Pub: Buros Institute
Title(s): Tests in Print
Language: English
Contact: Buros Institute of Mental Measurements, University of Nebraska, Lincoln NE 68588, USA
Govt/Priv: priv
Lang/Ed: ed
StanDef: test

Comments:

(see also record US01)

Objective/purpose:
A comprehensive index to MMYs (US03) published to date (the most recent edition located by the TFTS is 1983).

Target group:
Same as for MMY.

Procedure:
Includes any test appearing in MMY and actually in print and available for purchase or use.

Scope of Influence:
A standard reference.

Summary:
Covers the same scope as MMY. Entries contain publication/ordering information and a list of references to research and/or reviews of each test. Tests are classified by title, subject, and publisher.

Comments:
Like MMY, is very broad in scope, and so can be out of date. A good first reference for those beginning a review of the literature on published tests.

==========

US05: U.S.A.: "Questions to Ask When Evaluating Tests."

Record: US05
TFTSmem: Alderson
Country: U.S.A.
Auth/Pub: L. M. Rudner
Title(s): Questions to Ask When Evaluating Tests. ERIC/AE Digest, Department of Education, The Catholic University of America, Department of Education, O'Boyle Hall, Washington DC 20064. Document EDO-TM-94-06. April 1994
Language: English
Contact: Professor Caroline Gipps, Dean of Research, Institute of Education, University of London, 20 Bedford Way, London WC1H 0AL
Govt/Priv: priv
Lang/Ed: ed
StanDef: guideline

Comments:

A popularisation in simple form of the APA Standards. Useful and readable. 2 pages.

==========

US06: U.S.A.: "Criteria for Evaluation of Student Assess..."

Record: US06
TFTSmem: Alderson
Country: USA
Auth/Pub: National Forum on Assessment
Title(s): Criteria for Evaluation of Student Assessment Systems. Educational Measurement: Issues and Practice, Spring 1992
Language: English

Contact: Professor Caroline Gipps, Dean of Research, Institute of Education, University of London, 20 Bedford Way, London WC1H 0AL. Probably also available from FairTest -- see US07.
Govt/Priv: priv
Lang/Ed: ed
StanDef: guideline

Comments:

A one page list of 8 criteria and a short paragraph on each, describing and justifying.

==========

US07: U.S.A.: "Principles and Indicators for Student Assess..."

Record: US07
TFTSmem: Davidson
Country: U.S.A.
Auth/Pub: National Forum on Assessment
Title(s): Principles and Indicators for Student Assessment Systems (in press) 30 pp.
Language: English
Contact: National Center for Fair and Open Testing (FairTest), 342 Broadway, Cambridge, MA 02139 USA, ph: 617-864-4810, fax: 617-497-2224
Govt/Priv: priv
Lang/Ed: ed
StanDef: guideline

Comments:

Objective:
This document offers seven principles of student assessment. Each principle is accompanied by a number of indicators of its operation. As stated in 'Purposes of the Principles' on pp.1-2: "The *Principles* are intended to help transform assessment systems and practices as part of wider school reform. Assessment should support and be integrated with changes in instruction and curriculum that improve learning." ... "Each *principle* in this document defines a broad goal; it provides context and guidance for developing or refining an important part of the overall assessment system. The *indicators* are lists of more precise statements for use in developing or evaluating the system and its parts." On p. 3 the authors state that the intended users of this document are a wide array of people, e.g. policymakers, teachers, administrators, teacher training institutions, advocacy groups, and researchers, among others. The authors also note that quite a lot of assessment is going on in U.S. schools, and they hope that this document will provide "a special focus on the impact of assessment on instruction and learning" (p. 2).

Intended audience/target group(s):
(See citation under 'Objective', immediately above.)

Procedure:
The document begins with a discussion of proposed foundations and conditions for effective schools. There are then the seven principles, each followed by a number of indicators by which each principle can be seen at operation, or not, as the case may be. See examples under 'Summary' below.

Scope of influence:
The National Forum on Assessment published an earlier document called "Criteria for the Evaluation of Student Assessment Systems" (see record US06).

According to an earlier draft of the US07 document, the "Criteria have been endorsed by nearly 100 organizations." Hence, there is already a base on which to build the present document. More generally, FairTest (a key player in the present document) is an influential privately-funded watchdog in U.S. testing. They regularly lobby for changes in testing laws, and their newsletter (The FairTest Examiner) is a barometer of social change in testing in the U.S. It is therefore safe to assume that within the enterprise of test protest and change in the U.S., the present document should come to have considerable influence.

Summary:
The Foundations chapter states that the document's authors "agree on the following beliefs:

1. All students deserve a strong opportunity to learn high-level content in and across subject areas.
2. Thinking is the most basic and important skill.
3. All students deserve an equitable opportunity to learn in a resource-rich, supportive school.
4. High achievement takes many forms.
5. Equity demands similiarity in the standards of learning for all students and in the instructional quality offered to each student, together with the opportunity to demonstrate learning in a variety of ways.
6. Family and community support is essential to student success." (p.4)

The authors believe there are "four conditions [that serve as] a foundation for schools to ensure successuful learning and support the assessment practices promoted by the *Principles*:

1. Schools organize to support the multiple learning needs and approaches of all their members.
2. Schools work to understand how learning takes place and what facilitates learning.

3. Schools establish clear statements of desired learning for all students and help all students achieve them.
4. All schools have equitable and adequate learning resources and classroom conditions, including capable teachers, rich curriculum, safe and hospitable buildings, equipment and materials, and essential support services." (pp. 4-5)

The Principles and their indicators then follow. Each principle and its indictators are explicated in two pages. Following are the focused statements of each principle appearing as a header to each explication. There is no table of contents, so the following list also serves that function:

"Principle 1: The primary purpose of assessment is to improve student learning.
Principle 2: Assessment for other purposes supports learning.
Principle 3: Assessment systems are fair to all students.
Principle 4: Professional collaboration and development improves assessment practices.
Principle 5: The broad community participates in assessment development.
Principle 6: Communication about assessment is regular and clear.
Principle 7: Assessment systems are regularly reviewed and improved" (pp. 4-19).

As noted above, indicators are given for each principle, for example, for Principle 7: "A continuing group has responsibility for monitoring the assessment review process." and "Cost-benefit analysis of the assessment system focuses on its effects on instruction and learning." (p.19). The indicators are too numerous to report here, and the above two examples are illustrative only.

The document closes with a glossary, bibliography, and list of resource organizations. There will also be a two-page summary of the document issued separately for informational purposes.

Commentary:
An important document authored by and endorsed by the pro-active reformist test change movement in the USA.

==========

US08: U.S.A.: "Implementing Perform. Assessments: A Guide..."

Record: US08
TFTSmem: Davidson
Country: U.S.A.
Auth/Pub: Neill, Monty, Phyllis Bursh, Bob Schaeffer, Carolyn Thall, Marilyn Yohe and Pamela Zappardino. Published by FairTest. Cost (Summer 1995): US$ 6.00.
Title(s): Implementing Performance Assessments: A Guide to

Classroom, School and System Reform (No date is given, however I believe it was first published in 1994, based on a FairTest announcement of its arrival.) (57 pp.; includes a full-page FairTest membership and publication order form)

Language: English

Contact: National Center for Fair and Open Testing (FairTest), 342 Broadway, Cambridge, MA 02139 USA, ph: 617-864-4810, fax: 617-497-2224

Govt/Priv: priv

Lang/Ed: ed

StanDef: guideline

Comments:

Objective:
On p. 1, The document defines 'performance assessment' to include "things like classroom observation, projects, portfolios, performance exams and essays. These methods provide instructionally useful information by evaluating students on whether they can understand and can use knowledge, not just recognize, repeat and fill in bubbles." On the back cover, the authors indicate that this document is intended to be "a concise guide for teachers, administrators and others interested in using performance assessments in their classrooms and school systems."

Intended audience/target group(s):
(See second citation under 'Objective', immediately above.)

Procedure:
The document is a mixture of presentation formats: exposition, examples (of performance assessment), humor, and reports from news media on assessment reform, particularly as it involves performance assessment.

Scope of influence:
Compared to US07, this document seems aimed at a larger audience and should have a wider applicability. At the same time, it is more
focused on a single (albeit complex) thread in US testing reform: direct testing, which here is called performance assessment.

Summary:
Following is the Table of Contents listing of each chapter title:

"I. Introduction: New Assessments for Better Education
II. Classroom Assessment
III. Classroom Evaluation and Scoring
IV. Validity in Classroom Assessment and Evaluation
V. School-level Assessment and Evaluation
VI. Accountability
VII. Organizing for Change: What You Can Do
Resources
References"

Commentary:
As with US08, this document is an expression of ideals and ideas from the test reform movements in the USA. It is far more readable than US08, primarily because of its use of multiple presentation formats as noted above. By virtue of its focus on performance assessment (admittedly a broad term), it is a more narrow document than US08. However, it is clear that FairTest sees performance assessment as a valid, identifiable alternate paradigm to norm-referenced psychometric testing.

==========

APPENDIX FOUR: Contact Notes

This Appendix contains further general comments and notes on persons contacted who did not supply material for the TFTSTB or from whom material unrelated to standards was received, in the judgment of the TFTS Member supplying the notes. The TFTS Member is given for each set of comments/notes below.

TFTSTB Records: Commentary/Notes by Charles Alderson

Summary of part of ILTA TFTS survey conducted by Alderson

Alderson contacted 81 addressees, of whom 48 (59%) provided information.

A further 21 replied, but said they had no relevant information (these included 4 countries that did not otherwise provide responses:
Lesotho,
Germany,
Sweden
South Africa).

12 contacts failed to respond at all, including 6 countries that failed to respond to the Survey at all:
Botswana,
Brazil,
Kenya,
Malaysia,
Malta,
Nigeria

The 48 respondents included contacts from the following countries:
China
Hong Kong
India
Ireland
Mauritius
Namibia
The Netherlands
New Zealand
Portugal
Seychelles
Singapore
Switzerland
Tanzania
Uganda
United Kingdom (including England and Scotland separately)

also
ALTE ('Europe')
UNESCO ('France')

From the 48 respondents, 24 (50%) appeared to define 'standards' as 'guidelines', 8 seemed to define standards as 'performance' and 7 as 'test' (which could also be categorised as 'performance', making 15 in all - 31%). 8 responses were classified as 'other'.

I should point out that the Language Testing Research Group at

Lancaster had already conducted a survey of testing practice by EFL/ESL examining bodies in the UK, in the early 1990s. The initial results were published in a paper by Alderson and Buck in 1993: "Standards in Language Testing: A Survey of the Practice of UK Examination Boards in EFL Testing". Language Testing, 10(2): 1-26. The full results of the final survey are appear in Alderson, Clapham and Wall (1995) Language Test Construction and Evaluation. Cambridge University Press.

Since we had already amassed considerable data on the UK scene in EFL, I concentrated on non-EFL sources and contacts for the UK part of the ILTA Survey, and must therefore refer ILTA colleagues to the above two publications to complete the picture with respect to the 'state of the art' in language testing standards.

The ILTA charge letter implies that ILTA was interested in collecting information about 'good testing practice' in different countries. To that end, it appears to me that the responses from the following contacts offer the most relevant 'guidance' for the development of standards (ILTA charge letter, July, 1993)

Hong Kong provided a number of very interesting, relevant and highly professional documents from the Hong Kong Examinations Authority (HK01 is an excellent document worthy of serious further study; as is HK04 and HK05: Statistics Used in Public Examinations in Hong Kong.

Mauritius is trying to mauritianise its examination system, and documents received make it clear that standards are likely to be developed for the establishment and monitoring of item banks, and new examinations systems (see MA01).

The IAEA provided a number of extremely useful documents (NE02): the international surveys that are conducted under the IAEA aegis appear to follow detailed guidelines and standards, and as IAEA lists 53 member countries, it clearly has considerable experience in dealing with different expectations about standards and measurement. Of particular relevance are the IAEA Guidebook, and the Handbook (edited by Keeves) which is an editing of existing material "in a way that would report the experiences of IAEA research workers and would provide standards and guidelines, as well as advance appropriate methodology for future IAEA studies." The document is far too long to summarise but should be consulted in detail if ILTA decides to draw up its own Standards and Guidelines.

New Zealand provided a very useful and thorough document at least partially addressing ILTA concerns in the form of the Regulations and Prescriptions Handbook (NZ01).

In the UK the most important documents provided are: The GCSE Mandatory Code of Practice, 1993 (UK05) and The GCE A and

AS Code of Practice, 1994 (Record number UK19). Both documents are intended to promote quality and consistency in the examining process across all examining boards... help to ensure that grading standards are consistent in each subject and from year to year.. and provides for the development of a system that seeks to ensure consistency, accuracy and fairness in the operation of all .. examinations. "The documents are divided into 8 sections as follows:

1. Responsibilities of examining boards and examining board personnel
2. Syllabuses
3. Setting of question papers and provisional mark schemes
4. Standardisation of marking
5. Coursework assessment and moderation
6. Grading and awarding
7. The quality of language
8. Examining boards' relationships with centres.

The comment in the Database entry says it all: "This is exactly the sort of dccument TFTS was hoping to find and should be closely consulted in any discussion of ILTA standards"

The British Psychological Society (UK10) has a Steering Committee on Test Standards and issues guidelines on test standards, as well as certificates of competence in (occupational) testing, which covers subjects like defining assessment needs, basic principles of scaling and standardisation, reliability and validity, deciding when tests should be used, and much more: ILTA should certainly consult this and related documents (including Code of Conduct and Ethical Principles and Guidelines: UK13).

In addition, the UNESCO response (FR03) reports on the Monitoring Education-for-All-Goals Project. This is interesting to ILTA because it represents an international attempt to reach agreement on the qualities of measuring instruments used in their Survey, an attempt which included workshops on survey techniques (implying common standards), and which asserts the importance of avoiding cultural bias in measurement instruments and their development. ILTA will find it valuable to obtain further reports as they become available.

ALTE have developed a Code Of Practice based on the JCTP Code, which is very germane to ILTA's interests, and which should be looked at in some detail (EU01).

Finally, after this report was completed, Alderson received advance notice of a publication due in April entitled The Guide to Best Practice, following the launch of the British National Language Standards (UK08), which claims to "focus on assessment procedures and techniques drawing on examples from employers and training providers who are already

putting the language standards into practice". We also understand that work is underway to develop "National Cultural Standards", and "Standards for Interpreters and Translators". Obviously the development of standards, whatever they are defined as, is actively underway in relevant organisations in the UK.

Following are notes from particular correspondence:

E M Sebatane, Director, Institute of Education, National University of Lesotho, and President of IAEA: "In Lesotho we do not have any set of educational assessment standards as defined in your letter. Among the reasons for this are that educational assessment as a discipline is relatively new in the country, and that virtually no standardized tests are used....Test constructors in the case of public examinations are usually subject specialists who base their work mainly on the curriculum with no formal guidance on following technical qualities of the tests".

H. G Macintosh, Secretary/Treasurer of IAEA (International Association for Educational Assessment): "While I will be happy to help I am afraid that I do not have any relevant information for you but will obviously keep my eyes open"

Frances Ottobre, Executive Secretary, IAEA: Informed us of the AERA/APA/NCME committee to revise standards for educational and psychological tests, a chair of which is Eva Baker, and of Ron Hambleton's committee to "develop instrument adaptation standards" (which may be the committee that John de Jong is on). Also mentioned IAEA's cooperation with the latter committee by "sharing a copy of a report on its project 'International Test of Developed Abilities1 (ITDA), which was a study to determine the feasibility of developing a test of academic aptitude in four languages. The procedures used in the feasibility study followed what we believe are important guidelines for the construction of educational assessment instruments in general, and for developing tests in several languages in particular".

Caroline Gipps, Institute of Education, University of London: "I cannot at the moment think of anything appropriate for your request, but I will bring it up at the next meeting of ICRA, and we will get back to you if there is anything to report". Later response enclosed a copy of the ETS Code of Fair Testing Practices (see record US02), the ERIC/AE Digest EDO-TM-94-06 (US05) and the Criteria for Evaluation of Student Assessment Systems (US06), developed by National Forum on Assessment, but did not report anything from the UK.

Gunther Trost, Institut fur Test und Begabungsforschung, Bonn, Germany: "I am afraid I cannot offer you any information you do not already have". However, he mentioned Ebel (1965), Helmstadter (1966), Macintosh and Frith (1982), a chapter in

the International Encyclopaedia of Education on construction or selection of educational assessment instruments, ETS Guidelines, the AERA/APA/NCME revisions and The International Test Commission, whose President is Ron Hambleton.

Patricia Broadfoot, School of Education, University of Bristol: "I have asked my colleagues who are more closely involved in the field of language testing than I am myself, but have drawn a blank". Interestingly she enclosed a copy of promotional materials for the IELTS! The response does, however, suggest that our original request has been misinterpreted as pertaining exclusively to language testing.

Thomas Kellaghan, Educational Research Centre, St Patrick's College, Dublin: "No documents", but referred to the revision of test standards by AERA/APA/NCME.

Afzal Ahmed, The Mathematics Centre, West Sussex Institute of Higher Education, Bognor Regis: "No specific guidance material for the construction of assessment instruments."

John de Jong, CITO, Arnhem, Netherlands: "CITO has no official Institutional Instrument with respect to Standards. As a member of ALTE, however, we adhere to the ALTE Code of Practice."

Dr S P Kulshrestha, Institute for Studies in Psychological Testing, 101 Doon Vihar, Jakhan, Rajpur Road, Dehradun, India, responded that "we do not have any such standards or related information regarding TFTS in languages."

D J Barrett, University of Cambridge Local Examinations Syndicate, 1 Hills Road, Cambridge CB1 2EU referred the request to the Research and Evaluation Division, who, despite follow up letters from Alderson and Barrett, have not replied.

Ingemar Wedman, Professor of Educational Measurement, replied on behalf of H Mattsson, Division of Educational Measurement, Department of Education, University of Umea, S-90187, Sweden. "We do not have any specific standards but are using the standards developed by AERA, NCME and APA in the US, at least in general terms. Recently we have obtained the more specific standards developed by ETS, a document we presently are examining."

J Edmundson, Joint Council for the GCSE, 23-29 Marsh Street, Bristol BS1 4BP. The Joint Council coordinates the work of the Examining Groups on "policies that they have in common, but has not published any documentation for advice regarding assessment in languages though the GCSE Examining Groups may do so for the guidance of their syllabus committees/ examiners/ teachers. Modern Languages in the GCSE examinations are operated under the National Subject-Specific Criteria and will be governed in the future by the National Curriculum

Subject Orders for Modern Foreign Languages". Enclosed a copy of the National Criteria for GCSE French, which simply give assessment objectives and definitions of levels of attainment, and do not therefore represent standards as TFTS understands them. They are not included in the textbase at present.

Prof L N Verma, Industrial and Vocational Training Board, Sir Rampersad Neerunjun Complex, Ebene, Rose Hill, Mauritius, had no information about standards in language testing.

Peter Dickson, (Department of Professional and Curriculum Studies, National Foundation for Educational Research in England and Wales, The Mere, Upton Park, Slough, Berkshire SL1 2DQ, and also International Coordinator of the IAEA Language Education Study) suspected that he could only draw our attention to sources familiar to us. However, he copied our request to the first meeting of the IAEA Language Education Study national research coordinators in November 1994. We have received no response, but since the Current Steering Committee for the Study includes Alister Cumming (OISE) and Elana Shohamy, perhaps further information can be gleaned from them?

Dr P H Bredenkamp (South African Certification Council, P O Box 74299, Lynnwood Ridge 0040 South Africa) informed us that the Certification Council is not an examining body, so that they do not have their own guidelines to assessment or other documents. He referred us to the Department of National Education.

Donald McIntyre (British Educational Research Association University of Oxford, Department of Educational Studies 15 Norham Gardens, Oxford OX2 6PY England Tel 01865 274021) did not reply, but passed on our request to the Scottish Council for Research in Education, which did supply some documents (qv). The interesting inference is that BERA does not have its own guidelines for standards of assessment or testing practices.

Martin Taylor (Research Officer, Associated Examining Board, Stag Hill House, Guildford, Surrey, GU2 5XJ) referred us to SCAA and the Codes of Practice for GCSE and GCE A and AS Level (qv). He added: "I am not sure whether we have anything that will meet your needs. The GCE examining boards and GCSE examining groups maintain standards by means of a comparative process year-on-year: for example, in any examination grade A is awarded to work of the same quality as work which received grade A in the corresponding examination in the previous year." However, he gives no details of how this is achieved, and he clearly interprets standards to mean standards of performance or level of attainment..

Professor T Christie (Centre for Formative Assessment Studies, School of Education, University of Manchester, Oxford Road, Manchester M13 9PL) says that he is conscious of a 'gap in the market here' despite the wealth of practical experience in the

UK. He suggested that the reports of the three agencies involved in National Testing in England and Wales (NFER, CATs and STAIR) were all germane and were published by SEAC (the precursor to SCAA) in 1990. We have written for these but failed to obtain copies. Professor Christie adds: "Last week we ran a conference 'Establishing comparabilities: issues in setting standards for learning and instruction'. Several papers were much to the point. They will be published by Falmer Press as a book edited by myself and Bill Boyle in 1995."

R Chamberlain (University of Namibia, Private Bag 13301, 13 Storch Street Windhoek, Namibia. Fax (061) 307 2444) says that testing in Namibia is undeveloped, but that the South African Metric for school leavers is being replaced by the British UCLES' IGCSE examination. There appear to be no national standards at present, although a seminar is due to be held in 1995 to discuss the provision of valid tests for learners between school and further education or work.

A Oberholzer (Natal Education Department, Private Bag XI, Berea Road 4007, Durban, South Africa) provided a long description, in a letter, of the system of assessment and examination in Natal, where there are official syllabi, but only one formal examination. Subject departments have considerable freedom in assessment content and practice, although she reports a call for more standardisation and prescription of methods across subject departments in the Ministry. She mentions the South African Certification Council which was set up to ensure comparability of standards (see their response above), and the Independent Examinations Board (see records SA01 and SA02). She concludes by pointing out this is a time of major change and upheaval in South Africa, and things are likely to change rapidly.

Colleagues contacted, but with no response:

P Moanakwena, Research and Testing Centre, PO Box 189, Gaborone, Botswana

A Serpa de Oliveira, Fundacao, Cesgranrio, Rua Cosme Velho, 155, Brazil

Prof W B Dockrell, The University of Newcastle upon Tyne, Newcastle upon Tyne, NE1 7RU

A Yussufu, Kenya National Examinations Council, P O Box 73598, Nairobi, Kenya

A A Abdul Rahman, Malaysia Examinations Council, 13th and 14th Floor, KWSP Building, JLN Raja Laut, 50604 Kuala Lumpur, Malaysia

B M A Abdul Rahman, Malaysian Examinations Syndicate, Jalan

Duta, 50605 Kuala Lumpur, Malaysia

A Sammut, Test Construction Unit, Education Department, Floriana CMR 02, Malta

Prof D S J Mkandawire, Faculty of Education, University of Namibia, Private Bag 13301, Windhoek, Namibia

Prof Dr W H F W Wijnen, University of Limburg, Dept of Educational Development and Research, P O Box 616, 6200 MD Maastricht, Netherlands

Prof I Aina, National Business and Technical Examinations Board, PMB 1747, Benin-City, Nigeria

C Talbot, Independent Examinations Board, P O Box 875, Highlands North 2037 Johannesburg, South Africa

N N Mutanekelwa, Examinations Council of Zambia, P O Box 50432, Lusaka, Zambia

========== [end Alderson, Appendix Four] ==========

TFTSTB Records: Commentary/Notes by Ari Huhta

My task was to cover all European countries except Britain and Ireland. Usually, I contacted potential sources of testing standards by sending a letter. Sometimes I knew a person to whom to address the letter, but usually this was not the case. The response rate varied considerably; especially many Eastern European countries failed to send any answer whatsoever. This may indicate that there are no written guidelines on testing in many European countries.

The documents received covered both school-related achievement tests and independent proficiency tests for study or work purposes. The documents were usually in the national languages of the country where they originated from (English, German, French, Swedish, Finnish). Also, the languages tested by the examinations referred to in the documents varied: some related to tests of one language only (e.g. the French DELF / DALF exams), others referred to whole systems covering several languages (e.g. National certificate in Finland, International Baccalaureate, Corsortium for the European Certificate of Attainment in Modern Languages).

Most standards documents received concerned existing tests. Usually, they were guidelines on how to administer the tests in practice (how to set up exam room, advice for invigilators, etc) (e.g. CCSE, Goethe-Institut, and Finnish exams). Some documents were guidelines on test design and item writing (e.g. France and Finland), or guidelines for assessors and markers to standardise their work (e.g. Goethe-Institute and France).

Typically, a standards document had only one main purpose, such as guiding item writing, but on occasion it could cover more than one purpose, e.g. guiding test administration plus marking. It was fairly typical to see specialised documents which also gave more general advice on good testing practice.

An interesting observation that I made when reading the documents was that some guidelines used very similar content although they were designed to serve very different purposes. For example, the Goethe-Institute and Cambridge (CCSE) documents both present extensive examples of candidates' performance, accompanied with assessors' comments. For the Goethe-Institut, these are to guide assessors' work, whereas the Cambridge documents are intended for test takers and their teachers as examples and guidelines on what to expect in the examination, in order to help the teachers prepare their students for the test, and to facilitate students' decisions about the most appropriate examination level for them. Also, some sections in the French DELF/DALF guidebooks contain examples of test taker performance, but there the aim of the documents are to guide test design and item writing.

Another observation made during analysing the documents was the use of video tapes to complement written standards documents (e.g. CCSE, Finnish exams). Usually these contain benchmark examples of candidates' speaking ability for oral assessors to standardise their work. The CCSE exam also has a video tape that illustrates some practical considerations when administering the oral test.

Finally, I would like to comment briefly on some testing practices and contexts that appear to be fairly typical in many European countries, and that are related to the existence (or lack of) explicit written guidelines on any of the phases in test design and use. Typically the Scandinavian countries, but also many other European countries, have had very few large-scale testing systems. The number of test takers may have been large in some school-related examinations, but the exam systems have been fairly simple, with only a limited number of subjects or languages tested, and with a limited number of people involved in producing the tests. A good example is a school-leaving exam which is used only in one country and designed by a special, fairly small examinations board. The members of such boards usually serve for several years, and at any one time there are very few new members in the board. This means that most of the detailed information is passed on orally, or in letters and other brief, unofficial documents, among the members of the board. The members meet regularly, year after year, and there is little need to write e.g. test specifications on paper. Thus, one explanation for the scarcity of European guidelines on language testing is the non-complex nature of the examination systems in many countries. The only people in such centralised systems who need written guidelines are the administrators and invigilators of the exams (often teachers), and the test takers, of course. For

these groups, there appear to be more documents, which are usually very brief and often rather general in nature.

Examination systems which are complex -- i.e., they employ a greater number of people, they have several different tests for different purposes, and they administer the tests abroad -- need much more documentation in order for the system to function properly. Very good examples of such complex systems are the University of Cambridge Local Examinations Syndicate and the German Goethe-Institute. If a country, or a testing board, moves towards a more complex examination system, then there is a growing need for more written guidelines to keep the system together, and, indeed, such a trend can be seen for example in my home country, Finland.

Below is the list of persons and institutions that I contacted but who either failed to respond at all or who responded by sending something else instead of standards. As you notice, it was very hard to get any answers from institutions when there was no specific person to turn to.

I noticed that Charles had been able to 'extract' an answer from one or two people whom I also contacted but who failed to give me any response. I've left these people out from my list because, obviously, they responded to somebody in our group. I haven't thoroughly checked if there are any such instances left on my list, but keep this possibility in mind when putting together our information. Charles has such a wide network of personal acquaintances and friends that this overlap in our 'target audience' was unavoidable.

Persons contacted without response:

Alexander A. Barchenkov, Moscow Linguistic University (Russia)

J. Fisiak, Faculty of Modern Languages and Literature, Uniwersytet im Adama

Mickiewicza w Poznaniu (Poland)

Tony Fitzpatrick, Paedagogische Arbeitsstelle des Deutschen

Volkshochschul-verbandes (DVV) (Germany)

Einar Gudmunsson, University of Iceland (Iceland)

H. Hart, University of Utrecht, Faculty of Social Sciences (Holland)

Hristo Kaftandjiev, University of Sofia (Bulgaria)

Jean Max Kaufmann, Ministère d'Education Nationale (France)

Gabriella Pavan De Gregorio, CEDE (Italy)

Diana Rumpite, Foreign Languages Department, Riga Technical University (Latvia)

Eleonora Schmid, Hohere technische Bundeslehranstalt, Wien (Austria)

Institutions contacted without response:

Centre for Educational Research and Innovation, c/o OECD, Paris (France)

Centro de Informacoes e Reducoes Publicas - CIREP, Lisboa (Portugal)

Centro de Investigacion y Documentacion Educativa (CIDE), Madrid (Spain)

Comenius University of Bratislava, Faculty of Philosophy (Slovakia)

Comparative Education Society in Europe, Brussels (Belgium)

Ethniko kai kapodistriako panepistimio Athinon, Dept. of Education / Languages (Greece)

Generalitat de Catalunya, Direccio General de Politica Linguistica, Barcelona (Spain)

Hungarian Institute for Educational Research, Budapest (Hungary)

Hungarian Psychological Association, Budapest (Hungary)

Institut national de reserche pedagogique, Paris (France)

Institute of National Problems in Education, Moscow (Russia)

Kharkov State University, Faculty of Foreign Languages (Ukraine)

Ministère de l'Éducation nationale (France)

Ministère de l'Éducation nationale (Luxembourg)

Ministère della Pubblica Istruzione (Italy)

Ministerie van Onderwijs, Bestuur van het Hoger Onderwijs en het Wetenschappelijk Onderzoek (Belgium)

Ministry of National Education and Religious Cults, Section C 'Eurydice' (Greece)

Nizhnii Novgorod State Pedagogical Institute of Foreign Languages (Russia)

Pedagogical Research Institute, Kiev (Ukraine)

Pedagogical Research Institute, Minsk (Belarus)

Pyatigorsk State Pedagogical Institute of Foreign Languages (Russia)

Subdireccion General de Cooperacion Internacional, Madrid (Spain)

Univeridade de Lisboa, Departamento de Linguae Cultura Portuguesa (Portugal)

Universita degli Studi di Milano, Department of Modern Languages (Italy)

Universita degli Studi di Roma, Department of Modern Languages (Italy)

Universitá per Stranieri, Perugia (Italy)

Universitat zu Koln, Faculty of Education (Germany)

Universitat Wien, Faculty of Philosophy / Education (Austria)

Universitatea Bucuresti, Faculty of Psychology, Education and Sociology (Romania)

Universite de Geneve, Faculty of Psychology and Educational Sciences (Switzerland)

Universite de Paris - VII, Department of Modern Languages / Education (France)

University of Amsterdam, Faculty of Educational Science (Holland)

University of Ljubljana, Faculty of Arts and Sciences (Slovenia)

University of Warsaw, Faculty of Pedagogy (Poland)

University of Zagreb, Faculty of Education (Croatia)

Univerzita Karlova, Faculty of Philosoply (Chech Republic)

Vilnius University, Faculty of Philology (Lithuania)

World Association for Educational Research, Ghent (Belgium)

Persons and institutions who replied but did not send any material considered 'standards' documents:

Alliance Française, Paris (France) replied by saying that we should turn to the DELF / DALF commission for the information on standards.

Angela Hasselgren, Universitetet i Bergen (Norway) replied by sending information about the level of attainment in foreign languages in Norwegian schools.

M. Diachov, Institute for Ethnic Problems of Education, Moscow (Russia) replied by sending material on Russian language and literature curriculum for secondary schools.

Rüdiger Grotjahn, Ruhr-Universitat Bochum (Germany) replied by giving some names who might know about the existence of such documents in Germany.

Valmar Kokkota, Dept. of Foreign Languages, Tallinn Technical University (Estonia) replied that they don't really have written standards in use in Estonia.

Rainer H. Lehmann , Universität Hamburg , Fachbereich Erziehungswissenschaft (Germany) replied that there are not many document published in Germany that could be called standards, not at least in mother tongue education.

Madeline Lutjeharms, Vrije Universiteit Brussel (Belgium) replied that they don't have such material because they don't have a central language testing body in Belgium; only some guidelines to teachers in very general terms

========== [end Huhta, Appendix Four] ==========

TFTSTB Records: Commentary/Notes by Carolyn Turner

Regarding records for Canada and South Africa:

1. The areas that I covered were Canada and South Africa. Rather than an extensive coverage of an area of the world, I set out to cover 2 specific countries in an intensive manner. My findings were much more success in Canada than in South Africa.

OVERVIEW
In total, there were 20 documents received from Canadian sources and 3 from South African sources. In Canada, 18 of the documents were specifically related to language education evaluation (mainly received from provincial ministries of education), while the remaining 2 were general educational evaluation and psychological human measurement. In South Africa, 2 documents were specific to language education evaluation and 1 was general education standards.

To use the TFTS Report's categories, 13 of the Canadian documents were interpreted as guidelines for evaluation practices, 2 were performance standards, and 5 were tests. One of the records labeled as test was actually a bibliography of French tests used across Canada. One of the South African documents was interpreted as performance standards, 1 was a test, and 1 was "other".

METHODOLOGY
The methodology I used in Canada turned out to be very different from what I used in South Africa.

Methodology in Canada: (1) TFTS letter sent to all provincial ministries of education, federal language service branches, university language instruction divisions, agencies and associations such as the Canadian Psychological Association, and personal contacts in the field of language testing and evaluation across Canada; (2) #1 was followed up through telephone calls and e-mail; (3) for non-responders, a second round of letters was sent.

Methodology in South Africa: (1) TFTS letter sent to educational and private agencies; (2) #1 was 100% unsuccessful, so worked through a colleague Bonny Peirce from OISE who was temporarily in South Africa.

2. Examples and more detail on both countries

CANADA

a. Background: Canada is a bilingual country with French and English being the two official languages. The English-speaking provinces have French as an official second language, while Québec the 1 French-speaking province has English as the official second language. Federal documents are bilingual.

b. 2 Canada-wide documents referring to standards of practice were received (not language evaluation specific).

(1) PRINCIPLES FOR FAIR STUDENT ASSESSMENT PRACTICES FOR EDUCATION IN CANADA (1993): This is an exciting document in that its objective is to provide a nation-wide consensus on a set of principles and related guidelines generally accepted by professional organizations as indicative of fair assessment practice within the Canadian educational context (i.e., it strives for the fair and equitable assessment of all students.) Its 2 parts include: assessment carried out by teachers (elem, sec, post-sec) and standardized assessments developed external to the classroom by commercial test publishers, provincial and territorial ministries of education, etc.

Should be consulted by ILTA as a nation-wide effort for standards of practice for assessment.

(2) GUIDELINES FOR EDUCATIONAL AND PSYCHOLOGICAL TESTING (1986) - published by the Canadian Psychological Association (CPA): The objective is to provide criteria for the evaluation of tests, testing practices, and the effects of test use within Canada. Intended to be consistent with the APA Standards (USA), but due to the differing legal and social contexts (including the bilingual nature of Canada) content is Canadian specific. Used by the Public Service Commission of Canada: Language Division for test development of bilingual competency tests for all civil servants.

c. The majority of documents received were language-specific public education provincial documents referring to standards of practice. Public education standards of practice for language assessment and evaluation (if they exist) are set at the provincial level by the Provincial Ministry of Education.

(1) The English-speaking provinces mainly sent specific FSL tests being used in their school systems (see RECORD CA03 - ANNOTATED LIST OF FRENCH TESTS, 1992).

(2) The French-speaking province (Québec) send many documents pertaining to guidelines in practice specifically related to language education. In Québec, 2 school systems exist side by side and therefore the province is preoccupied with L2 instruction (both ESL and FSL). It provides much guidance to its educators in terms of evaluation procedures and test construction. For example in June 1993, the Québec Ministry of Education (MEQ) changed the provincial secondary

leaving exam concerning ESL oral proficiency from individual interviews to group testing. Preceding this change, documents and workshops were provided for educators. (See RECORD CA14, GUIDE D'ÉVALUATION DE LA PRODUCTION D'UN DISCOURS ORAL EN CLASSE: ANGLAIS, LANGUE SECONDE, 2e CYCLE DU SECONDAIRE)

d. University documents received were mainly specific language tests and accompanying guidelines (e.g. OTESL, Ontario Test of ESL originally used with nonnative speakers of English entering post-secondary educational institutions; CanTEST/TESTCan, used with nonnative French or English speakers wanting to benefit from either university study or professional exchange in Canada.)

There are many adult language programs going on in universities across Canada. Only one document was received, however, which included assessment/evaluation guidelines for teachers or administrators of these programs: ENGLISH LANGUAGE PROGRAM: INTENSIVE CURRICULUM, U. of Alberta, Extension Centre. (See RECORD CA19).

e. It was interesting to learn what university professors/researchers consult when they are asked to construct a government test (e.g. MEPT, Midwives English Proficiency Test, David Mendelsohn and Gail Stewart at York University, Ont.). They use a variety of sources besides their own expertise. In communicating with professors across Canada, I found there was a pattern in their sources. (See EXPLANATORY NOTES REGARDING RECORDS FROM CANADA just before record CA01 in the TFTSTB for a list of references.)

3. South Africa

a. Background - Due to the political transition in South Africa, standards for language evaluation mainly appear to focus on standards of student performance rather than standards of practice for test development. Much work is beginning to take place at the university level.

b. Few documents received: one dealt with standards of practice: HANDBOOK FOR ENGLISH: GEC EXAMINATION (1994) - the International Exam Board (IEB). At present, the IEB which is a non-profit organization provides curriculum design and exams for schools and adult education in many subjects including language.

The information provided above is a general outline of documents received from 2 countries in the TFTS Report, Canada and South Africa. Following are notes on persons or agencies contacted but from whom I received no response.

Persons:

Roger Mareschal. Public Service Commission of Canada, Language Services Directorate. Asticou Centre, Cit des Jeunes Boulevard, Hull, Québec, Canada K1A 0M7

Gerard Monfils. Public Service Commission of Canada, Evaluation Consultant, Language Training Branch (LTPB), Tests, Measurement and Evaluation Service (TMES), 15 Bisson St., Hull, Québec, Canada K1A 0M7

Wally Lazaruk. Alberta Education. Language Services Branch. 11160 Jasper Ave., Edmonton, Alberta, Canada T5K 0L2

Institutions and Associations:

Association for the Study of Evaluation in Education in South Africa (ASEESA). South Africa

Institute of Psychological Research, Inc. (Human Resources Testing). 34 Fleury St. West, Montréal, Québec, Canada

========== [end Turner, Appendix Four] ==========

TFTSTB Records: Commentary/Notes by Elaine Wylie

1. Introduction: I focused on Australia. One of the documents I reviewed was published in the UK, but relates to the testing system IELTS (International English Testing System) which is a joint UK/Australia project.

2. Method: I used a combination of standard letters and personal approaches by phone and e-mail. Like my colleagues I found personal approaches much more effective. Letters went to all universities and boards of secondary school studies in all states and territories.

University responses were very poor. Several responded that they had nothing relevant, and only my own, where I made personal contact, came up with any documents. Boards of senior secondary school studies were better. Of the eight states and territories in Australia, four sent documents, and there was also a document from NAFLaSSL (the National Assessment Framework for Languages at Senior Secondary Level), a scheme which facilitates the offering of languages which otherwise would have too small a candidature for any one state to handle (e.g. Khmer and Ukrainian).

In a number of cases my personal contacts were to ask permission to include documents which I already had in my possession, e.g. the "IELTS Specifications".

In particular, in relation to personal contacts, I would like to thank Susan Zammit, from the Australian Council for Educational Research (ACER), who spent a great deal of time going through archives, and came up with what I believe is the oldest document in the collection, Record No. AU05 "Educational objectives being tested in the Commonwealth Secondary Scholarship Examination" (1967). Another ACER document worth noting is Record AU01, "Ethical considerations in educational testing" by Geoff Masters. It succinctly sets out responsibilities for ACER test developers and for users of ACER tests.

3. The Oz angle: I have been trying to decide what is distinctive about the Australian collection. It is probably the case that, because in Australia there is a strong emphasis on school-based assessment, even for school-leaving purposes, many of the documents are guidelines to help teachers devise assessment systems and activities. There is one set of documents which I believe provides a very good model of communication between a central policy-making body and teachers. These are Record No. AU12, Discussion papers published by the Assessment Unit of the Board of Secondary School Studies of Queensland, my own state, which has been in the forefront of school-based assessment. Those by Royce Sadler in particular have had

considerable impact on educational assessment in other states. They are concerned with education in general.

Record AU17, ESL Development: Language and Literacy in Schools is a very interesting set of documents about school-based language assessment. The project from which the publication arose was undertaken by the National Languages and Literacy Institute of Australia recently, and co-ordinated by Penny McKay. Essentially it involved the development of three sets of Bandscales for school students at three age levels - Junior Primary, Middle/Upper Primary and Secondary levels of schooling, three sets of exemplar assessment activities to guide teachers in developing classroom activities which will provide data appropriate for data for relating students" language abilities to the Bandscales, and guidelines for using both the tasks and for reporting. The Bandscales are being used in a number of states for diagnostic, research and other purposes, and we are hoping that they may be used as a basis for allocating and distributing federal and state funds for ESL programs.

4. Conclusion: To conclude, I would like to quote from Record UK22, the IELTS specifications, developed by the IELTS team at UCLES, University of Cambridge Local Examinations Syndicate, in consultation with an international advisory committee. In the Code of Practice, it specifically gives a commitment to "Supporting the activities of professional associations involved in developing and implementing professional standards and codes, making available the results of research, and seeking peer review of its activities", a philosophy which I feel ILTA would support in whatever is the next stage of the standards exercise.

Notes:

Australian institutions which replied but did not send information:

Curtin University of Technology (West Australia) sent information on their matriculation requirements.

The University of Ballarat (Victoria) International Affairs Office replied that they were unable to assist with any materials.

University of Canberra (ACT) replied that they do not have their own standards.

Australian institutions contacted without response:

Australian Capital Territory Board of Studies
Department of Education, Northern Territory
Senior Secondary Assessment Board of South Australia
Victorian Board of Studies

The University of Adelaide
Australian Catholic University
The Australian National University
Bond University
University of Central Queensland
Charles Sturt University
Deakin University
Edith Cowan University
The Flinders University of South Australia
James Cook University
La Trobe University
Macquarie University
The University of Melbourne
Monash University
Murdoch University
The University of Newcastle
The University of New England
The University of New South Wales
Northern Territory University
The University of Queensland
Queensland University of Technology
University of South Australia
University of Southern Queensland
The University of Sydney
University of Tasmania
University of Technology, Sydney
Victoria University of Technology
The University of Western Australia
The University of Western Sydney
The University of Wollongong

APPENDIX FIVE: Chapters One and 11 of Alderson, Clapham and Wall (1995) Language Test Construction and Evaluation (Cambridge U.P.)
[Per arrangement with the publisher, available in the first one hundred printed reports only, and not in the ERIC Document or any other copies of this report.]

Note: the page numbers at the bottom of the following pages follows the pagination of this report. The photocopies also include the original pagination from the Alderson book.